Construction of a T7 phage random peptide library by combining seamless cloning with *in vitro* translation

(seamless cloning と *in vitro* translation の組み合わせによる T7 ファージランダムペプチドライブラリーの構築)

東　克暁

主論文

Regular paper

Biotechnology

Construction of a T7 phage random peptide library by combining seamless cloning with *in vitro* translation

Katsuaki Higashi, Sakiho Oda, Mai Fujii, Fumiya Nishida, Hayato Matsumoto, Jyoji Morise, Shogo Oka and Motohiro Nonaka*

Department of Biological Chemistry, Human Health Sciences, Graduate School of Medicine, Kyoto University, 53 kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan.

Running title

A new method for T7 phage library construction

***Corresponding author.**

Motohiro Nonaka, Ph.D.

18 Department of Biological Chemistry, Human Health Sciences, Graduate School of Medicine,

19 53 kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan.

20 TEL: +81-75-751-3954

21 E-mail address: nonaka.motohiro.4r@kyoto-u.ac.jp

22

23

Abstract

T7 phage libraries displaying random peptides are powerful tools for screening peptide sequences that bind to various target molecules. The T7 phage system has the advantage of less biased peptide distribution compared to the M13 phage system. However, the construction of T7 phage DNA is challenging due to its long 36 kb linear DNA. Furthermore, the diversity of the libraries depends strongly on the efficiency of commercially available packaging extracts. To address these issues, we examined the combination of seamless cloning with cell-free translation systems. Seamless cloning technologies have been widely used to construct short circular plasmid DNA, and several recent studies showed that cell-free translation can achieve more diverse phage packaging. In this study, we combined these techniques to construct four libraries (CX7C, CX9C, CX11C, and CX13C) with different random regions lengths. The libraries thus obtained all showed diversity $> 10^9$ plaque forming units (pfu). Evaluating our libraries with an anti-FLAG monoclonal antibody yielded the correct epitope sequence. The results indicate that our libraries are useful for screening peptide epitopes against antibodies. These findings suggest that our system can efficiently construct T7 phage libraries with greater diversity than previous systems.

41 *Keywords:*

42 Phage library

43 Peptide screening

44 Diversity

45 Seamless cloning

46 *In vitro* translation

47 **Introduction**

48

49 T7 phage display technology has been used to screen peptide ligands for various target
50 molecules. The T7 phage grows at a high rate; plaques can usually be detected within 2–3 h,
51 which is faster than the growth and plaque formation process of the filamentous M13 phage
52 (8–12 h)(1). In addition, peptide libraries constructed on T7 phage scaffolds exhibit less amino
53 acid sequence bias than libraries based on the M13 phage(2). These technical differences arise
54 mainly from differences in the morphogenesis and process of the lifecycle of the two phages.
55 After assembly in the periplasm, constituents of the M13 phage must be secreted from the cell
56 membrane. Thus, only peptides compatible with this pathway are displayed on the phage,
57 resulting in a loss of library diversity. T7 phages, in contrast, are assembled in the bacterial
58 cytoplasm and released upon host cell lysis, which does not generally limit peptide expression.
59 Given these advantages, the T7 phage has been frequently selected to screen short random
60 peptides *in vitro*(3–5) and *in vivo*(6–8).

61 One drawback to T7 phage display technology is the difficulty of library construction. The
62 T7 phage's genomic DNA consists of 36 kb of long linear double-stranded DNA (dsDNA). In
63 contrast to short circular plasmid DNA, the purification of T7 phage genomic DNA and

64 subsequent preparation of vector arms is a complex process requiring multiple steps. To
65 compensate for these difficulties, a cloning kit is currently available from Novagen (Merck
66 Millipore)(9). This cloning system includes restriction enzyme digestion and ligation
67 reactions(6, 10). The insert DNA can be prepared by polymerase chain reaction (PCR) using
68 primers with restriction enzyme sites at both ends, after which cohesive ends are obtained using
69 restriction enzymes. The insertion of random oligonucleotides using a restriction enzyme and
70 ligase raises several concerns. In addition to the insertion of unwanted amino acid sequences
71 directly after the g10 protein sequence, restriction enzyme sites in random sequences can also
72 cause unintended digestion, resulting in a loss of diversity. The ligation step involves a three-
73 fragment reaction, in which the target DNA is inserted between the left and right arms of the
74 phage DNA. A previous report demonstrated that the efficiency of phage packaging with the
75 cloning kit is around 10^8 plaque forming units (pfu)/µg (the µg represents mass of phage
76 DNA)(6).

77 Seamless cloning is a powerful tool in molecular biology, as it leaves no unwanted
78 nucleotides between the insert and the vector. This technique is useful when multiple inserts
79 need to be cloned in a specific order or reading frames must be conserved. In recent years,
80 homology-based cloning technologies such as Gibson Assembly(11), In-Fusion cloning(12)

81 and NEBuilder HiFi DNA Assembly(13), which is an improved Gibson Assembly product,

82 have facilitated seamless cloning. The principle of Gibson Assembly and NEBuilder is based

83 on three enzymes: T5 exonuclease, Phusion DNA polymerase, and Taq DNA ligase. T5

84 exonuclease degrades DNA ends in the 5′ to 3′ direction, yielding single-stranded DNA

85 overhangs. After an annealing step between the insert and vector overhangs, Phusion DNA

86 polymerase fills in the complementary chain and then Taq DNA ligase seals the gap. Therefore,

87 seamless cloning is advantageous over conventional ligation methods in that target genes can

88 be cloned at any position, regardless of the presence of restriction enzyme sites. However, this

89 method has not been rarely applied to long linear DNA such as T7 phage genomic DNA, and

90 it has not been applied at all to phage libraries. For this application, the possibility of T5

91 exonuclease affecting the terminal DNA sequence of the T7 phage should be considered, but

92 no reports have yet verified it.

93 Another concern regarding the preparation of a T7 phage library is related to the phage

94 packaging step. To date, preparation of the T7 phage packaging extract remains a hurdle, and

95 this step generally depends on the kit from Novagen (Merck Millipore) described above(9).

96 Meanwhile, an *in vitro* translation system has reportedly enabled the packaging of more than

97 10^{10} monoclonal T7 phages(14–16) and the rebooting of newly designed phages(17, 18). In

98 this system, the T7 RNA polymerase encoded in the phage genomic DNA is first expressed in
99 the kit extract, which triggers the expression of a series of phage components. Interestingly,
100 this *in vitro* reaction allows for not only the biosynthesis and assembly of phage proteins but
101 also the packaging of phage genomic DNA into capsids. To our knowledge, the application of
102 this principle to library construction has not been reported, though previous studies have
103 applied it to prepare designed monoclonal phages. Our new technology provides a
104 breakthrough in the T7 phage system that overcomes the disadvantage of inefficiency in
105 conventional methods of constructing phage libraries.

106 In this study, we developed a simple three-step process to generate high-diversity T7 phage
107 libraries displaying short cyclic peptides (CX7C, CX9C, CX11C, and CX13C). First, using T7
108 phage genomic DNA as a template, PCR was performed for regions 20 kb upstream and 16 kb
109 downstream of the target. Next, the resulting two-segment PCR products were connected using
110 the NEBuilder HiFi DNA assembly system. Finally, phages were synthesized and packaged
111 using an *in vitro* translation system. Verification using the generated libraries confirmed the
112 presence of intact sequences at both ends of the T7 genomic DNA sequence. The performance
113 of library construction was validated by screening against an anti-FLAG M2 monoclonal
114 antibody. Our system provides a platform for facile construction of T7 phage libraries with

115 high diversity.

116

117 **Materials and methods**

118

119 *Strains and culture conditions of the bacteria and bacteriophages*

120 *Escherichia coli* strain BLT5403 [F^-, *ompT*, *hsdS*$_B$ ($r_B^- m_B^-$), *gal*, *dcm* pAR5403 (Amp^R)]

121 and bacteriophage T7Select 10-3b were obtained from Merck Millipore (Darmstadt, Germany).

122 BLT5403 cells were cultured in lysogeny broth (LB) medium (Nacalai Tesque, Inc., Kyoto,

123 Japan) containing 100 μg/ml of ampicillin sodium salt (LB-Amp; Nacalai Tesque, Inc.).

124

125 *Phage titration*

126 Phage libraries or phage pools were diluted in LB medium or SM buffer (50 mM Tris-HCl,

127 pH 7.5, 100 mM NaCl, 10 mM $MgSO_4$, and 0.01% gelatin) at appropriate dilution rates.

128 BLT5403 cells in the exponential growth phase and diluted phages were added to 3–4 mL of

129 LB top agar (0.75%) and plated on Petri dishes. The plates were incubated at 37°C and the

130 plaques were counted.

131

132 *Construction of peptide-displaying phage libraries*

133 Plasmid pAR5403 was isolated from BLT5403 cells using a FastGene Plasmid Mini Kit
134 (Nippon Genetics Co., Ltd., Tokyo, Japan) according to the manufacturer's instructions.
135 T7Select 10-3b DNA (Merck Millipore) was digested with *Hin*dIII (New England Biolabs,
136 Beverly, MA, USA). The left phage arm was prepared through PCR using 10 ng of digested
137 phage genomic DNA as the template, 30 pmol of the forward primer 10-3b_1-34 (5′-
138 TCTCACAGTGTACGGACCTAAAGTTCCCCCATAG-3′), and 30 pmol of the reverse
139 primer 10-3b_20396-20421_Rv (5′-AGAATTCGGATCCCCGAGCATCACAC-3′) in a final
140 volume of 100 μl. The right phage arm was constructed using forward primers containing
141 random oligonucleotides CX7C-fw, CX9C-fw, CX11C-fw, and CX13C-fw (5′-
142 TGTGATGCTCGGGGATCCGAATTCTGGAAGCGGTTCTGGTTCATGT[NNK]$_n$TGCTA
143 AGCTTGCGGCCGCACTCGAGTAAC-3′, n = 21, 27, 33, 39, respectively, where K
144 represents equimolar amounts of T and G, and N is equimolar amounts of A, T, G, and C), and
145 the reverse primer 10-3b_36214-36249_Rv (5′-
146 AGGGACACAGAGAGACACTCAAGGTAACACCCAAAG-3′). These two reactions were
147 performed using KOD One PCR Master Mix (Toyobo Co., Ltd., Osaka, Japan) according to
148 the manufacturer's instructions. The PCR products were purified through phenol/chloroform

149 extraction and ethanol precipitation. Briefly, 3 M sodium acetate and distilled H_2O were added

150 to the reaction solutions to obtain 300 µl of 0.3 M sodium acetate solution, and an equal volume

151 of phenol/chloroform/isoamyl alcohol (25:24:1) (Nacalai Tesque, Inc.) solution was added. The

152 PCR products were precipitated with the addition of 750 µl ethanol and incubation for 2 h at

153 −30°C, followed by centrifugation for 30 min at 4°C and 16,200 × *g*. The precipitates were

154 washed with 1 ml of 70% ethanol, air-dried, and dissolved in TE (10 mM Tris-HCl, pH 8.0,

155 and 1 mM ethylenediaminetetraacetic acid [EDTA]). The purified left and right arms (0.3 pmol

156 each) were assembled using NEBuilder HiFi DNA Assembly Master Mix (New England

157 Biolabs) in a total volume of 60 µl according to the manufacturer's instructions. Then, 100 ng

158 of the assembled phage DNA and 36 ng of plasmid pAR5403 were added to 9 µl of LS70

159 Master Mix, a component of the myTXTL T7 Expression Kit (Daicel Arbor Biosciences, Ann

160 Arbor, MI, USA), to reach a final volume of 12 µl and incubated at 29°C for 16 h. The titers

161 of the prepared phage libraries were counted as described above. The value of phage packaging

162 efficiency [pfu/µg] is calculated by dividing the number of total phage plaques (pfu) by the

163 amount of DNA used (µg). The phages were then amplified through growth in 500 ml of liquid

164 medium with BLT5403 until lysis occurred. The amplified libraries were centrifuged at 4°C

165 and 12000 × *g* for 30 min and then the supernatants were filtered using a cell culture filter unit

166 (polyethersulfone, 0.22 µm) with a 500-ml capacity (Thermo Fisher Scientific, Waltham, MA,

167 USA). Polyethylene glycol (PEG) 8000 and NaCl were added to the filtrates at final

168 concentrations of 10% and 1 M, respectively, and the mixture was incubated on ice overnight

169 to precipitate phages. The PEG-precipitated solutions were centrifuged at 4°C and 12000 × *g*

170 for 1 h and then the supernatants were discarded. The pellets were resuspended in 60 ml of SM

171 buffer and dispensed into 1.6-ml aliquots. The libraries were stored at −80°C.

172

173 *Biopanning using the anti-FLAG antibody M2 clone*

174 The wells of an Immobilizer Amino F8 Module Clear Kit (Thermo Fisher Scientific) were

175 coated with 100 µl of 1 µg/ml of monoclonal anti-FLAG M2 antibody (Sigma-Aldrich, St.

176 Louis, MO, USA) diluted with coating buffer (30 mM Na_2CO_3 and 70 mM $NaHCO_3$, pH 9.6)

177 at 4°C overnight. The wells were washed three times with phosphate-buffered saline (PBS) and

178 blocked with 10 mM ethanolamine diluted in coating buffer at room temperature for 1 h. After

179 the blocked wells were washed with PBST (PBS-containing 0.05% Tween20) three times, 200

180 µl of phage library mixture containing 32.5 µl from each of four phage libraries along with 3%

181 bovine serum albumin and 0.1% Tween20 were added. After 30 min of binding, the wells were

182 washed ten times with PBST (except for the 4th and 5th rounds, which used 0.25 M and 0.5 M

183 NaCl in PBST, respectively) and then the phages were eluted with 50 µl of 1% sodium dodecyl
184 sulfate (SDS) in PBS. The eluted phages were collected in 10 ml of LB-Amp medium, and the
185 titer was measured by taking a portion of eluted phages, diluting it appropriately, and then
186 plating it. The relative binding efficiency was calculated by dividing the number of phages
187 retrieved from the well after washing by the number of phages added to the well. The eluted
188 phages in LB medium were amplified through incubation with BLT5403 cells at 37°C until
189 lysis occurred. The subsequent rounds of panning were performed using the same methods as
190 in the 1st round.
191
192 *Sample preparation for next-generation sequencing (NGS)*
193 For the preparation of NGS samples, the 1st and 2nd PCR steps were performed using
194 PrimeStar HS DNA Polymerase (Takara Bio Inc., Shiga, Japan). The 1st PCR was performed
195 using phage libraries and screened phage pools as templates. For primers, NGS_1st_F (5′ -
196 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGTCAGGTGTGATGCTCGG-3′)
197 and NGS_2nd_R (5′ -
198 GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCTGCGTTAGCGTCACCTT-
199 3′) were used as the forward and reverse primers, respectively. The resulting 1st PCR products

200 were purified using a FastGene Gel/PCR Extraction Kit (Nippon Genetics Co., Ltd.) according
201 to the manufacturer's instructions. Next, the 2nd PCR was performed using the primer set from
202 the Nextera XT Index Kit (Illumina, San Diego, CA, USA) with the 1st PCR products as
203 templates. The 2nd PCR products were analyzed via gel electrophoresis using 1.5% agarose in
204 TAE (Tris, acetic acid, and EDTA) buffer, and the desired bands were extracted. The products
205 were purified using a FastGene Gel/PCR Extraction Kit (Nippon Genetics Co., Ltd.). The
206 sequences of the purified DNA products were analyzed with the Illumina iSeq 100 sequencing
207 system (Illumina).
208
209 *NGS data analysis*
210 The Fastq files generated by the iSeq 100 system were analyzed using Python 3.9.12 scripts.
211 After translation of the sequence reads into amino acid sequences, the peptides displayed
212 following the sequence of the g10 capsid protein were extracted. The number of reads per
213 sequence was counted and normalized so that the total number of reads equaled 10^5. Finally,
214 cyclic sequences (e.g., sequences ending with cysteine residues) were extracted. To evaluate
215 the diversity of each library, Shannon's diversity index was calculated as described
216 previously(19, 20). Then, the diversity index J', known as Pielou's evenness index, was

217 calculated by dividing the logarithm of Shannon's diversity index values by the logarithm of
218 the number of samples(21). Pielou's evenness index ranges from 0 to 1 as diversity varies from
219 low to high. For WebLogo 3(22) analysis in the screening using the anti-FLAG M2 antibody,
220 the amino acid residues at positions of −1, +1, +2, and +3 from the YK motif were extracted
221 and the top 100 reads were subjected to analysis. Graphs were drawn using GraphPad Prism
222 software.

223

224 *PCR of phage clones in the packed libraries for verification of DNA ends*

225 Four amplified libraries were mixed and plated on petri dish at appropriate dilution rates
226 and plates were incubated at 37°C until plaques appeared. 8 clones were picked up and
227 amplified. PCR was performed using PrimeStar HS DNA Polymerase (Takara Bio Inc.) with
228 the following primers designed to amplify the ends of the phage genomic DNA; 1-15 (5' -
229 TCTCACAGTGTACGG -3'), 1-18 (5' - TCTCACAGTGTACGGACC -3'), 6-23 (5' -
230 CAGTGTACGGACCTAAAG -3'), 14-29 (5' - GGACCTAAAGTTCCCC -3'), 181-200_Rv
231 (5' - AGGTGACTTTAGGAGGATAC -3'), 185-200_Rv (5' – AGGTGACTTTAGGAGG -3'),
232 36001-36022 (5' - TATCAGTGTCACCTAAAGAGGG -3'), 36001-36018 (5' -
233 TATCAGTGTCACCTAAAG -3'), 36234-36249_Rv (5' - AGGGACACAGAGAGAC -3'),

234 36230-36249_Rv (5' - AGGGACACAGAGAGACACTC -3'), 36236-36245_Rv (5' -
235 ACACAGAGAGACACTCAAGG -3'), 36212-36231_Rv (5' -
236 TCAAGGTAACACCCAAAGAC -3'). The reaction solutions were analyzed via gel
237 electrophoresis.
238
239 *Comparison of phage packaging efficiency of myTXTL and packaging extract*
240 For myTXTL reaction, 100 ng of intact phage genomic DNA, 1 nM of pAR5403 and 2%
241 of PEG5000 were added to 9μl of myTXTL master mix with a final volume of 12 μl. For
242 packaging extract, 100 ng of intact phage genomic DNA was added to 5 μl of packaging extract
243 with a final volume of 6 μl. The titers of phages were measured by taking a portion of the eluted
244 phages, diluting it appropriately, and then plating it. The packaging efficiencies were calculated
245 as described above.
246
247

248 **Results**

249

250 *Construction of T7 phage libraries*

251 One of the major difficulties in T7 phage library construction lies in the diversity of
252 libraries. First, we evaluated the efficiency of phage production per unit mass using phage
253 genomic DNA in the TXTL reaction and the packaging reaction with the packaging extract.
254 While previous report has indicated the production of up to approximately 10^{13} pfu of
255 monoclonal phages with addition of dNTPs into the reaction solution to replicate phage
256 genomic DNA(16), our goal is to generate phage libraries with independent sequences.
257 Therefore, we conducted the reactions without adding dNTPs to the reaction mixture in order
258 to prevent DNA replication. Using 100 ng of phage DNA in both experiments, the TXTL
259 reaction resulted in the production of 1.26×10^9 pfu (1.26×10^{10} pfu/μg), while 9.57×10^6 pfu
260 (9.57×10^7 pfu/μg) of phages were produced using the packaging extract, indicating a 132-
261 fold increase in efficiency (Fig. 1A).

262 We next aimed to prepare four phage libraries (CX7C, CX9C, CX11C, and CX13C).
263 To achieve DNA assembly using two linear segments for four libraries, we conducted long
264 PCR of the T7 phage genome regions 20 kb upstream and 16 kb downstream of the target (Fig.

265 1B). For each downstream PCR, the DNA sequence that overlaps the upstream 25 bases and a

266 sequence encoding random peptides were added to the 5′ end of the primer. The image obtained

267 from agarose electrophoresis showed the PCR products of both fragments as single bands (Fig.

268 1C). Next, the two obtained fragments were assembled using the NEBuilder HiFi DNA

269 assembly kit. The reactant samples showed new bands above the upstream and downstream

270 fragments, indicating that assembly was achieved (Fig. 1C, arrowhead). We then performed *in*

271 *vitro* translation and the packaging reaction using the myTXTL system. The resulting

272 diversities were 2.6×10^9 (CX7C), 2.6×10^9 (CX9C), 3.6×10^9 (CX11C), and 2.9×10^9

273 (CX13C) (Fig. 1C).

274

275 *Low sequence diversity bias in the T7 phage libraries*

276 The practical diversity of phage libraries is determined by the bias of random sequences in

277 the library and the efficiency of phage packaging. To evaluate the bias of the obtained libraries,

278 we performed NGS analysis. A comparison of the counts per sequence for each library showed

279 that out of hundreds of thousands of reads, a maximum of six duplicated clones was detected,

280 and the majority of sequences had only one count, indicating that little duplication occurred in

281 the amino acid sequences of our random peptide libraries (Fig. 2). Next, the evenness of the

clones in the prepared library was examined using Pielou's evenness index. This index has been used to assess the T-cell receptor repertoire(21). Pielou's J' values, which are calculated for each library by dividing the Shannon index H' by the maximum H' value (i.e., H_{max}), were > 0.998 for all libraries. This result indicates that our short random peptide libraries were highly unbiased.

Verification of the DNA ends of the clones in the packaged libraries

In the NEBuilder reaction, the 5′ ends of dsDNA are generally removed by T5 exonuclease. As T7 phage genomic DNA is linear, the terminal DNA is likely to be included, which may impair library performance. To assess the effect of exonuclease activity, the terminal DNA was amplified through PCR using primers targeting each of the two ends (Fig. 3A). In this analysis, if the terminal DNA is lost, product bands will not be observed. PCR using genomic DNA from the three clones showed that all DNA removed from the library was intact (Fig. 3B). The removed terminal DNA may have been repaired by an intrinsic mechanism in *E. coli* after the NEBuilder and packaging reactions. These results indicate that the activity of T5 exonuclease in the NEBuilder reaction solution does not affect the prepared libraries.

299 *Validation screening using an anti-FLAG monoclonal antibody*

300 To further test library performance, we screened the libraries against an anti-FLAG
301 monoclonal antibody, M2 clone. This clone was originally obtained by immunizing mice with
302 the FLAG peptide DYKDDDDK. The epitope of the resulting M2 antibody is DYKXXD(23).
303 An antibody-coated plate or control non-coated plate was incubated with a mixture of four
304 libraries and the bound phages were eluted and amplified in *E. coli.* We performed a total of
305 six rounds. The out/in ratio indicated a 715-fold increase from the 1st screen (2.39×10^{-5}) to
306 the 6th screen (1.71×10^{-2}), demonstrating that the phage was well enriched (Fig. 4). The ratio
307 of phage titers in antibody-coated plates to non-coated plates increased by 10,683-fold during
308 the screening process, further indicating that the phage was sufficiently enriched.
309

310 *NGS analysis of the screened phage pools*

311 Next, the displayed peptide sequences of the phage pool were analyzed using NGS. Notably,
312 the top 20 peptide sequences in the 6th round all contained the consensus sequence YK and
313 most exhibited the correct epitope sequence DYKXXD (Fig. 5A), indicating that our screening
314 system and the random peptide libraries were valid. There was one sequence
315 “CAVRESIDCYKC” showing 10 amino acids between two cysteines. It is possible that the

316 three-base shifted sequence was coincidentally included in the oligo DNA synthesis process.
317 For quantitative analysis of the obtained sequences, the frequency of occurrence of amino acid
318 sequences was visualized using WebLogo 3. The top 100 sequences containing the YK motif
319 in each round were extracted, and the amino acids at positions −1, +1, +2, and +3 from YK
320 were visualized. As a result, the DYKXXD sequence was clearly identified in the 2nd round
321 (Fig. 5B and C), indicating that two rounds of screening were sufficient to retrieve the FLAG
322 epitope sequence in our study. When the matching peptides were categorized by peptide length,
323 it was observed that 13-mer peptides showed a higher rate of matches, followed by 11-mer
324 peptides, in this screening process (Fig. 5D). This result suggests that the FLAG epitope was
325 originally a linear epitope and thus binds preferentially to peptides from longer libraries, which
326 have linear shapes.
327
328
329 **Discussion**
330
331 Phage random peptide libraries have been applied to a variety of screening applications.
332 Traditionally, they have been used to screen for epitope sequences of antibodies and binding

333 motifs of target molecules. However, their application is not limited to such analyses. For
334 example, screening with anti-carbohydrate (Le^a) antibodies led to the identification of a series
335 of epitope peptides that mimic the structure of carbohydrates(24, 25). D-peptides that bind to
336 natural target molecules can be obtained using the mirror-image screening technique(5, 26).
337 Furthermore, due to its robust structural stability, T7 phage can be screened *in vivo*(7, 8, 27).
338 Thus, the T7 system has the potential for future application to a wide variety of biological
339 investigations.

340 In general, the quality of a peptide library can be defined by high diversity and low bias in
341 the displayed sequences. As T7 phage lyses bacterial cells simultaneously with amplification,
342 the peptide sequences are not affected by the phage's life cycle, resulting in the production of
343 unbiased peptide libraries. However, the diversity of libraries has been limited to date due to
344 the reliance on complex genetic manipulation and packaging methods(6). In this study, we
345 established a novel approach to simplify the library construction process and to increase library
346 diversity. In the previous protocol, preparing the insert DNA involves 4-step experiments(28).
347 Firstly, randomized library DNA and reverse complementary strand extension primer are
348 annealed. The annealed duplex DNA is extended by Klenow fragment DNA polymerase I. Then,
349 the DNA is digested by two restriction enzymes. Finally, the prepared insert DNA is ligated

350 with double digested phage DNA by restriction enzymes (vector arm) for making the full length
351 of phage DNA. In contrast, the method described in this report simplifies the process to only
352 two steps. Researchers only need to perform PCR of the two fragments of the phage DNA and
353 then employ the NEBuilder reaction to connect the parts of the phage DNA. Moreover, the
354 TXTL system dramatically simplifies the library construction process, particularly when
355 considering per diversity size along with the increase of efficiency of reaction. This is because
356 of the high efficiency of producing phages by using TXTL system, which is over 100 times
357 compared to previous method of using packaging extract. With respect to amino acid bias, the
358 conventional method involving the use of two restriction enzymes for insert DNA digestion
359 leads to the loss of enzyme recognition sequences, introducing slight bias. However, our
360 method avoids such specific losses, resulting in lower bias compared to the previous approach.
361 Notably, no effect on the genomic DNA of the resulting library was observed from removal
362 of the terminal DNA by T5 exonuclease, as verified through PCR (Fig. 3). However, the repair
363 mechanism of the terminal DNA remains unclear. One possible mechanism is a T7 phage-
364 specific DNA replication system in the host *E. coli*(29). T7 phage DNA contains a 160-bp
365 terminal repeat sequence at both ends. During replication, intermediates containing overhangs
366 hybridize with each other to form head-to-tail concatemers. In the packaging step, the linked

367 DNA is cleaved by site-specific nucleases to release genomic monomers. In this study, the
368 reaction with NEBuilder may have removed approximately 50 bp of T7 phage DNA from both
369 termini, although hybridization could still occur with the remaining 60 bp of the repeat
370 sequence. In that case, the intact phage DNA would eventually be repaired. The performance
371 of the libraries was demonstrated through model screening using an anti-FLAG M2
372 monoclonal antibody, which has been tested in previous studies(23, 30, 31). In total, six rounds
373 of panning reactions were performed, and the epitope sequences were detected after the 2nd
374 round (Fig. 5B). Interestingly, the epitope sequence DYKXXD was most frequently detected
375 in peptides derived from CX13C libraries (Fig. 5D). Given that the FLAG epitope is a linear
376 epitope, longer peptides, which exhibit a greater degree of linearity than other cyclic peptides,
377 are more likely to provide matches. Overall, our system is useful for generating highly diverse
378 T7 phage libraries and can be applied to other visualization systems, including other antibodies
379 (e.g., Fab, Scfv, and VHH) and proteins, in the future.

380

381 **Author contributions**

382 Conceptualization, M.N.; Investigation, K.H., S.O., F.N., H.M., and M.F.; Software,
383 K.H. and M.N.; Supervision, J.M., S.O., and M.N.; Visualization, K.H.; Writing – original draft,

K.H. and M.N.; Writing – review and editing, J.M., S.O., and M.N.; Funding acquisition, M.N.

Disclosure of potential conflicts of interest

The authors have no potential conflicts of interest related to this work.

Declaration of competing interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the JST FOREST Program (grant number JPMJFR2167).

References

1. Green, M. R., and Sambrook, J. (2017) Plating bacteriophage M13. *Cold Spring Harb.*

401 *Protoc.* 2017, db.prot093427
402 2. Krumpe, L. R. H., Atkinson, A. J., Smythers, G. W., Kandel, A., Schumacher, K. M.,
403 McMahon, J. B., Makowski, L., and Mori, T. (2006) T7 lytic phage-displayed peptide
404 libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide
405 libraries. *Proteomics*. 6, 4210–4222
406 3. Hatanaka, T., Ohzono, S., Park, M., Sakamoto, K., Tsukamoto, S., Sugita, R., Ishitobi, H.,
407 Mori, T., Ito, O., Sorajo, K., Sugimura, K., Ham, S., and Ito, Y. (2012) Human IgA-binding
408 peptides selected from random peptide libraries: affinity maturation and application in
409 IgA purification. *J. Biol. Chem.* 287, 43126–43136
410 4. Yoneyama, T., Hatakeyama, S., Tobisawa, Y., Yamamoto, H., Imanishi, K., Okamoto, T.,
411 Tokui, N., Sugiyama, N., Suzuki, Y., Kudo, S., Yoneyama, T., Hashimoto, Y., Koie, T.,
412 Kamimura, N., Fukuda, M. N., and Ohyama, C. (2013) Blood group antigen-targeting
413 peptide suppresses anti-blood group antibody binding to antigen in renal glomerular
414 capillaries after ABO-incompatible blood reperfusion. *Transplantation*. 95, 418–425
415 5. Nonaka, M., Mabashi-Asazuma, H., Jarvis, D. L., Yamasaki, K., Akama, T. O., Nagaoka,
416 M., Sasai, T., Kimura-Takagi, I., Suwa, Y., Yaegashi, T., Huang, C.-T., Nishizawa-Harada,
417 C., and Fukuda, M. N. (2021) Development of an orally-administrable tumor vasculature-

418 targeting therapeutic using annexin A1-binding D-peptides. *PLoS One*. 16, e0241157

419 6. Fan, X., Venegas, R., Fey, R., van der Heyde, H., Bernard, M. A., Lazarides, E., and
420 Woods, C. M. (2007) An in vivo approach to structure activity relationship analysis of
421 peptide ligands. *Pharm. Res.* 24, 868–879

422 7. Sugihara, K., Kobayashi, Y., Suzuki, A., Tamura, N., Motamedchaboki, K., Huang, C.-T.,
423 Akama, T. O., Pecotte, J., Frost, P., Bauer, C., Jimenez, J. B., Jr, Nakayama, J., Aoki, D.,
424 and Fukuda, M. N. (2014) Development of pro-apoptotic peptides as potential therapy for
425 peritoneal endometriosis. *Nat. Commun.* 5, 4478

426 8. Pleiko, K., Põšnograjeva, K., Haugas, M., Paiste, P., Tobi, A., Kurm, K., Riekstina, U.,
427 and Teesalu, T. (2021) In vivo phage display: identification of organ-specific peptides
428 using deep sequencing and differential profiling across tissues. *Nucleic Acids Res.*
429 10.1093/nar/gkaa1279

430 9. Krumpe, L. R. H., and Mori, T. (2006) The use of phage-displayed peptide libraries to
431 develop tumor-targeting drugs. *Int. J. Pept. Res. Ther.* 12, 79–91

432 10. Fukunaga, K., and Taki, M. (2012) Practical tips for construction of custom Peptide
433 libraries and affinity selection by using commercially available phage display cloning
434 systems. *J. Nucleic Acids*. 2012, 295719

435 11. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith,
436 H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases.
437 *Nat. Methods*. 6, 343–345
438 12. Irwin, C. R., Farmer, A., Willer, D. O., and Evans, D. H. (2012) In-fusion® cloning with
439 vaccinia virus DNA polymerase. *Methods Mol. Biol.* 890, 23–35
440 13. Birla, B. S., and Chou, H.-H. (2015) Rational design of high-number dsDNA fragments
441 based on thermodynamics for the construction of full-length genes in a single reaction.
442 *PLoS One*. 10, e0145682
443 14. Rustad, M., Eastlund, A., Marshall, R., Jardine, P., and Noireaux, V. (2017) Synthesis of
444 infectious bacteriophages in an E. coli-based cell-free expression system. *J. Vis. Exp.*
445 10.3791/56144
446 15. Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016) The all E. coli TX-TL
447 toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol.* 5, 344–355
448 16. Garenne, D., Thompson, S., Brisson, A., Khakimzhan, A., and Noireaux, V. (2021) The
449 all-E. coliTXTL toolbox 3.0: new capabilities of a cell-free synthetic biology platform.
450 *Synth. Biol.* 6, ysab017
451 17. Mitsunaka, S., Yamazaki, K., Pramono, A. K., Ikeuchi, M., Kitao, T., Ohara, N., Kubori,

452 T., Nagai, H., and Ando, H. (2022) Synthetic engineering and biological containment of
453 bacteriophages. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2206739119
454 18. Pulkkinen, E. M., Hinkley, T. C., and Nugen, S. R. (2019) Utilizing in vitro DNA assembly
455 to engineer a synthetic T7 Nanoluc reporter phage for Escherichia coli detection. *Integr.*
456 *Biol.* 11, 63–68
457 19. Stewart, J. J., Lee, C. Y., Ibrahim, S., Watts, P., Shlomchik, M., Weigert, M., and Litwin,
458 S. (1997) A Shannon entropy analysis of immunoglobulin and T cell receptor. *Mol.*
459 *Immunol.* 34, 1067–1082
460 20. Held, H. A., and Sidhu, S. S. (2004) Comprehensive mutational analysis of the M13 major
461 coat protein: improved scaffolds for C-terminal phage display. *J. Mol. Biol.* 340, 587–597
462 21. Maceiras, A. R., Almeida, S. C. P., Mariotti-Ferrandiz, E., Chaara, W., Jebbawi, F., Six,
463 A., Hori, S., Klatzmann, D., Faro, J., and Graca, L. (2017) T follicular helper and T
464 follicular regulatory cells have different TCR specificity. *Nat. Commun.* 8, 15067
465 22. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence
466 logo generator. *Genome Res.* 14, 1188–1190
467 23. Slootstra, J. W., Kuperus, D., Plückthun, A., and Meloen, R. H. (1997) Identification of
468 new tag sequences with differential and selective recognition properties for the anti-FLAG

469 monoclonal antibodies M1, M2 and M5. *Mol. Divers.* 2, 156–164

470 24. Hatakeyama, S., Sugihara, K., Shibata, T. K., Nakayama, J., Akama, T. O., Tamura, N.,

471 Wong, S.-M., Bobkov, A. A., Takano, Y., Ohyama, C., Fukuda, M., and Fukuda, M. N.

472 (2011) Targeted drug delivery to tumor vasculature by a carbohydrate mimetic peptide.

473 *Proc. Natl. Acad. Sci. U. S. A.* 108, 19587–19592

474 25. Fukuda, M. N., Ohyama, C., Lowitz, K., Matsuo, O., Pasqualini, R., Ruoslahti, E., and

475 Fukuda, M. (2000) A peptide mimic of E-selectin ligand inhibits sialyl Lewis X-dependent

476 lung colonization of tumor cells. *Cancer Res.* 60, 450–456

477 26. Funke, S. A., and Willbold, D. (2009) Mirror image phage display--a method to generate

478 D-peptide ligands for use in diagnostic or therapeutical applications. *Mol. Biosyst.* 5, 783–

479 786

480 27. Mann, A. P., Scodeller, P., Hussain, S., Braun, G. B., Mölder, T., Toome, K., Ambasudhan,

481 R., Teesalu, T., Lipton, S. A., and Ruoslahti, E. (2017) Identification of a peptide

482 recognizing cerebrovascular changes in mouse models of Alzheimer's disease. *Nat.*

483 *Commun.* 8, 1403

484 28. Krumpe, L. R. H., and Mori, T. (2014) T7 lytic phage-displayed peptide libraries:

485 construction and diversity characterization in *Methods Mol. Biol.* (Nixon, A. E., ed.) Vol.

486 1088, pp. 51–66, Springer Science+Business Media, Berlin

487 29. Weigel, C., and Seitz, H. (2006) Bacteriophage replication modules. *FEMS Microbiol.*
488 *Rev.* 30, 321–381

489 30. Reyes, S. G., Kuruma, Y., Fujimi, M., Yamazaki, M., Eto, S., Nishikawa, S., Tamaki, S.,
490 Kobayashi, A., Mizuuchi, R., Rothschild, L., Ditzler, M., and Fujishima, K. (2021) PURE
491 mRNA display and cDNA display provide rapid detection of core epitope motif via high-
492 throughput sequencing. *Biotechnol. Bioeng.* 118, 1736–1749

493 31. Srila, W., and Yamabhai, M. (2013) Identification of amino acid residues responsible for
494 the binding to anti-FLAG™ M2 antibody using a phage display combinatorial peptide
495 library. *Appl. Biochem. Biotechnol.* 171, 583–589

496

497 *This is a pre-copyedited, author-produced version of an article accepted for publication in The*
498 *Journal of Biochemistry following peer review. The version of record of Higashi, K., Oda, S.,*
499 *Fujii, M., Nishida, F., Matsumoto, H., Morise, J., Oka, S., Nonaka, M. (2023) Construction of a*
500 *T7 random peptide library by combining seamless cloning with in vitro translation. J. Biochem.*
501 *is available online at: http://doi.org/10.1093/jb/mvad077.*

502

503 **Figure legends**

504

505 **Fig. 1.** Construction of short random peptide libraries using T7 phage. (**A**) Comparison of
506 phage packaging efficiency of myTXTL and packaging extract. (**B**) Overall schematic diagram
507 of library construction. (**C**) Agarose gel electrophoresis of DNA fragments after PCR and the
508 NEBuilder reaction. The *Hin*dIII digestion site is located in the multi-cloning site of T7 phage
509 10-3b DNA. For PCR products, representative upstream and downstream fragments of library
510 CX7C are shown. (**D**) Diversity of the resulting T7 phage libraries. pfu, plaque-forming units.

511

512 **Fig. 2.** Peptide sequence bias in each library. Sequence-by-sequence counts resulting from
513 NGS in the CX7C, CX9C, CXC11, and CX13C libraries. The total read counts and numbers
514 of clones in each library are shown. Shannon index H' and Pielou index J' values were
515 calculated as described in the Materials and methods section.

516

517 **Fig. 3.** Verification of DNA ends after library preparation. (**A**) Primer sets were adjusted for
518 length and designed to have equal melting temperatures. (**B**) The results of PCR verification.
519 T7 phage clones were selected from the prepared CX7C library and PCR was performed using

520 their DNA as the template. For each clone, validation was performed with four primer sets at

521 both DNA ends. The results presented are representative images of three clones.

522

523 **Fig. 4.** Model screening using the anti-FLAG monoclonal antibody, M2 clone, as the target

524 molecule. For the panning reaction, a mixture of the CX7C, CX9C, CX11C, and CX13C

525 libraries was used. Phage pools collected from antibody-coated wells (Ab[+]) were used in the

526 following round. Relative binding efficiencies were calculated as described in the Materials

527 and methods section.

528

529 **Fig. 5.** NGS analysis of model screening for anti-FLAG antibodies. (**A**) Sequencing results

530 sorted by NGS counts in the 6th round. DNA sequences in Fastq files were translated into

531 amino acids, and cyclic peptides with two cysteine residues at both ends were extracted and

532 ranked by count. (**B**) Sequence logo analysis. After extracting YK motif-containing amino

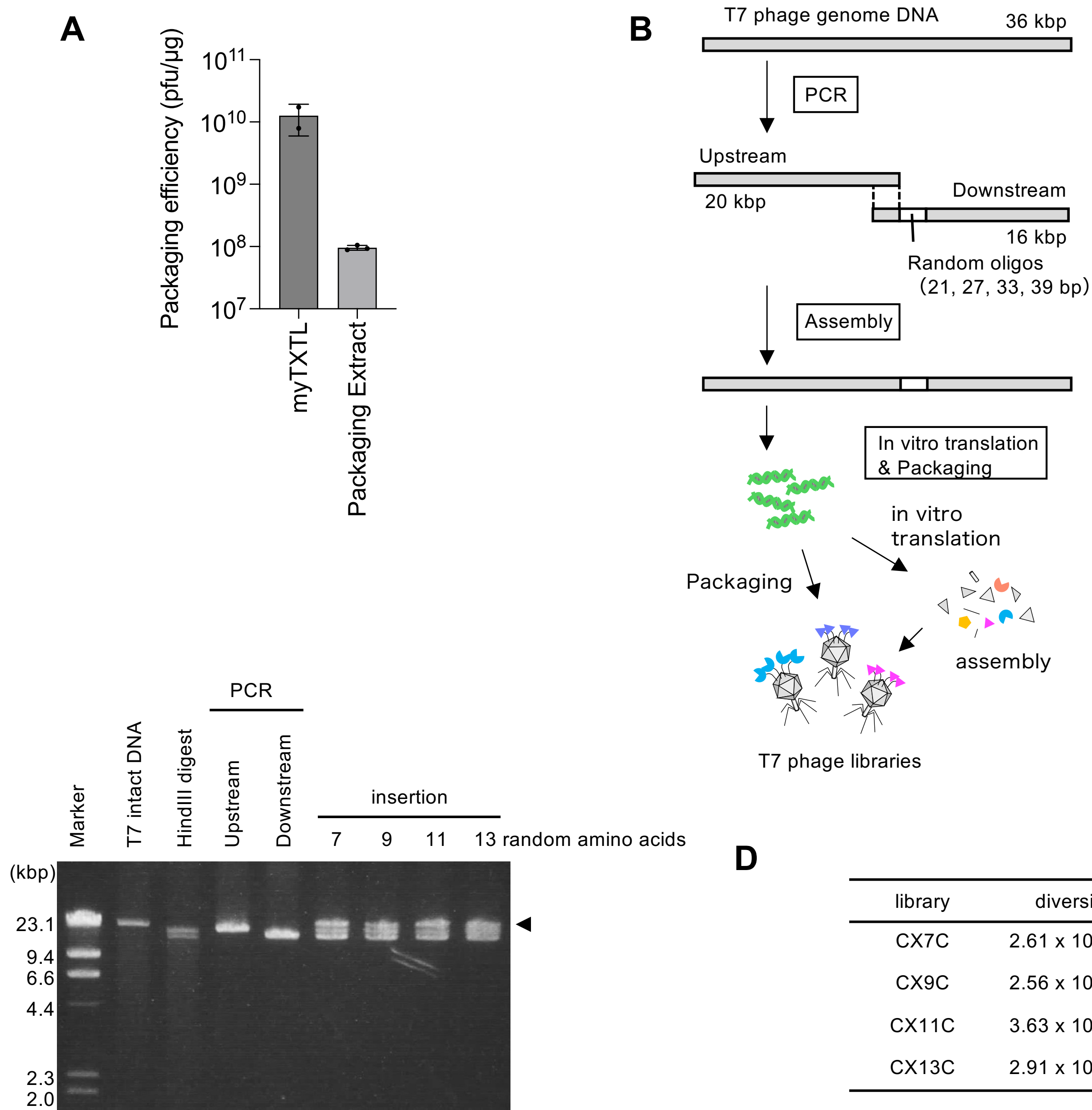
533 acids at positions −1 to +3, the top 100 sequences in terms of counts were extracted. The

534 sequences obtained in each round were analyzed using WebLogo3 and represented as bits (left)

535 and probability (right). (**C**) Number of YK sequences in the top 100 sequences. (**D**) Differences

536 in the occurrence of DYKXXD domains in the top 100 sequences of each of the four libraries.

Fig. 1



D

library	diversity
CX7C	2.61 x 10^9 pfu
CX9C	2.56 x 10^9 pfu
CX11C	3.63 x 10^9 pfu
CX13C	2.91 x 10^9 pfu

京都大学

Fig. 2

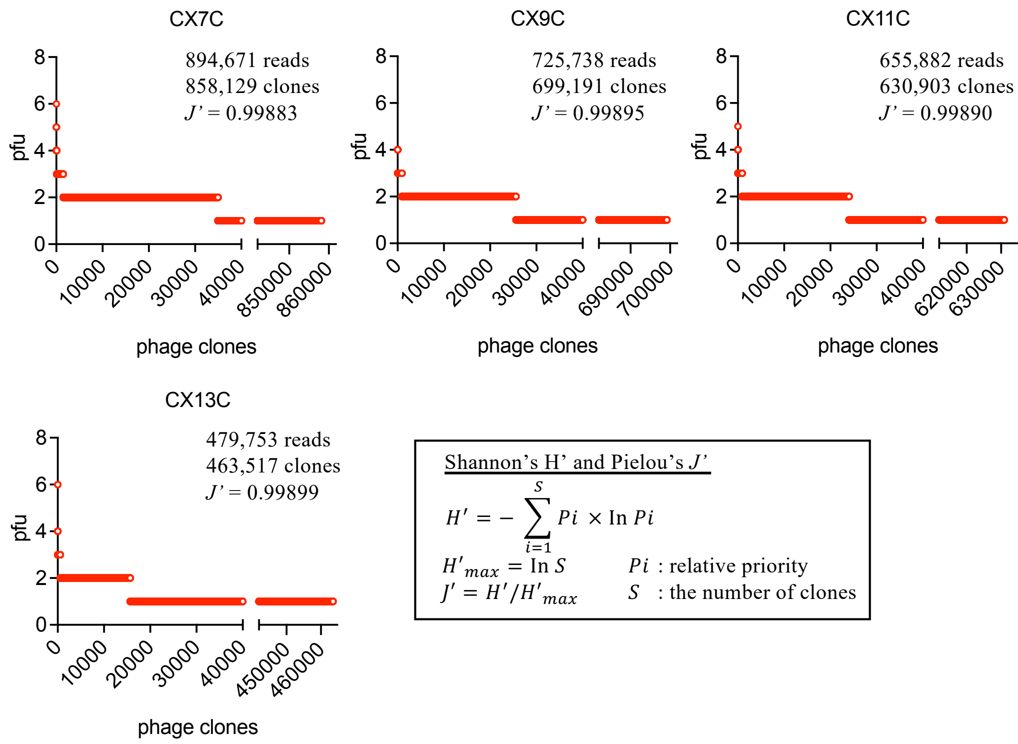


Fig. 3

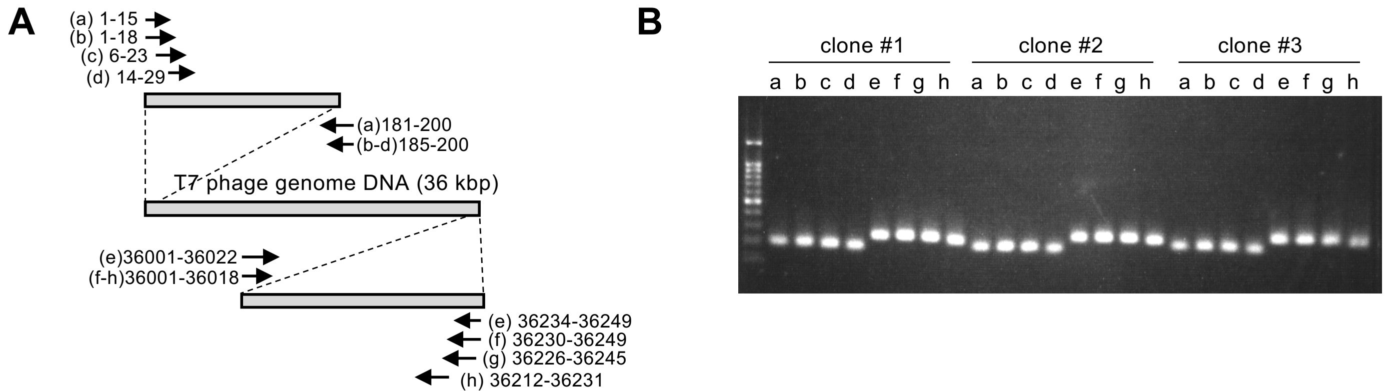
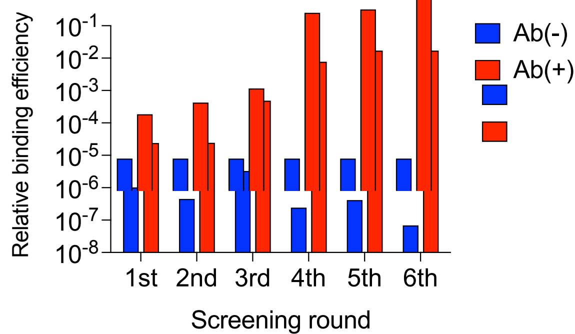


Fig. 4



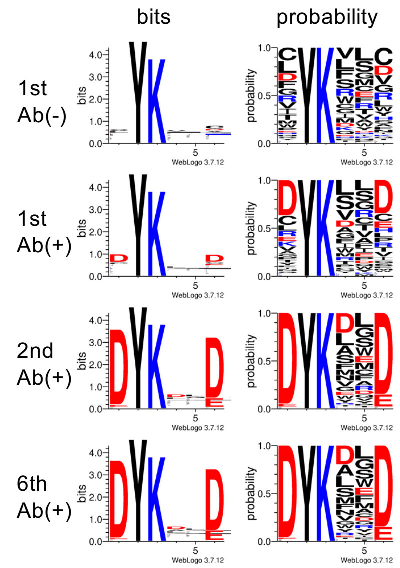
京都大学

Fig. 5

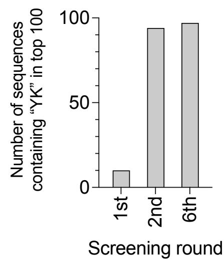
A

NGS count	Sequences	Number of random amino acids
6793	CLAVSRLLDYKDMDC	13
3398	CTRSPQGEDYKTYDC	13
2765	CAVRESIDCYKC	10
1977	CSGSVLDYKAVDC	11
1585	CFNVPAWLDYKGSDC	13
1409	CHRDYKLWDGLKC	11
1316	CGHTVWEYKDADC	11
1074	CTDGFMGDYKWTDGC	13
1057	CLLDYKMSESHLGRC	13
976	CAGCLADYKDEDGGC	13
939	CEDYKMGDRRGDNGC	13
855	CKWSRADDYKYLDKC	13
742	CAQDYKDRDWRCQAC	13
736	CAGFVGDRDYKASDC	13
726	CLCDYKWEESNGNSC	13
725	CPVSTSYVDYKSSDC	13
718	CRVSDNFFDYKVSDC	13
707	CADYKDCDERSYC	11
697	CTGDDYKNSDVGLWC	13
696	CSFSGRDYKAEDC	11

B



C



D