

1 **Efficient and flexible strategies for gene cloning and** 2 **vector construction using overlap PCR**

3

4 Zhongtian Liu¹, Tingting Zhang^{1, 2}, Kun Xu^{1*}

5 ¹College of Animal Science & Technology, Northwest A&F University, Yangling,
6 Shaanxi, 712100, China

7 ²Research Institute of Applied Biology, Shanxi University, Taiyuan, Shanxi, 030006,
8 China

9

10

11 *To whom correspondence should be addressed. Tel: +86-029-87092102; E-mail:
12 xukunas@nwsuaf.edu.cn.

13

14

15

16

17 **Abstract**

18 Gene cloning and vector construction are basic technologies in modern molecular
 19 biology for gene functional study. Here, we present flexible and efficient strategies for
 20 gene cloning and vector construction using overlap PCR. We firstly cloned the open
 21 reading frames (ORFs) of the porcine *MSTN*, chicken *OVA* and human *α -glucosidase*
 22 genes by overlap PCR-based assembling of their exons, which could be amplified
 23 with genomic DNAs as the templates without RNA extraction and RT-PCR reaction.
 24 Secondly, we generated additionally three designed functional cassettes by overlap
 25 PCR-based assembling of different DNA elements, which facilitated the construction
 26 their expression vectors greatly. Moreover, we further developed an interesting
 27 overlap-circled PCR method for fast plasmid vector construction without any cutting
 28 and ligating procedure. These advanced applications of overlap PCR provide useful
 29 alternative tools for gene cloning and vector construction.

30

31 **Keywords:** gene cloning; overlap PCR; exon assembly; circled PCR; vector
 32 construction

33

34 **Introduction**

35 Gene cloning and vector construction are the foundational and routine technologies
 36 for molecular biological research. The first step for studying a particular gene is
 37 usually to clone it from the cDNA using RT-PCR, and routinely the next step is to
 38 construct corresponding expression vector by restrictive endonuclease-based digestion
 39 and ligase-based ligation. The restrictive endonuclease-ligase dependent strategy for
 40 gene cloning has been deeply developed and widely used. However, there are
 41 inherently several limitations, such as the requirement of high-qualified cDNA
 42 template for desired gene cloning, the inconvenience for simultaneous multiple gene
 43 cloning and the limited restriction enzyme sites available on the backbone plasmid for
 44 vector construction. Multiple fragment cloning can be achieved by enzymatic
 45 assembly of overlapping DNA fragments with the alternative application of
 46 5'-exonucleases [1, 2] and by Golden Gate cloning using type IIs restrictive enzymes
 47 [3-5].

48 Besides, a series of nuclease or/and ligase-independent strategies for constructing
 49 vectors based on different characteristics of intent terminal sequences has been
 50 developed, such as the enzyme-free cloning using annealing tails [6], the T-A cloning
 51 applying special ends of PCR products[7], the UDG or USER cloning with Uracil
 52 DNA glycosylase [8, 9], the Gateway cloning dependent on bacterial homologous
 53 recombination (HR) [10], the mating-assisted genetically integrated cloning (MAGIC)
 54 using bacterial *in vivo* site-specific endonuclease cleavage and HR [11], and the
 55 sequence and ligation-independent cloning (SLIC) using *in vitro* HR and single-strand

annealing (SSA) [12]. Although these methods mentioned above all have their own applications and advantages, new improvements leading to more flexible, feasible, convenient and economical approaches for gene cloning and vector construction are still needed.

Here, we demonstrated that the open reading frame (ORFs) of functional genes could be assembled with exon amplicons from genomic DNA using overlap PCR strategy, which is useful and practical for amplifying long eukaryotic genes interrupted by introns [13-17]. Besides, multiple genes or DNA elements are sometimes required to be cloned into a single plasmid vector, which is always inconvenient with multiple cloning steps and confined by limited cloning sites available. The overlap PCR strategy could be effective for assembling these genes or elements. We further demonstrated that the overlap-circled PCR could be used to assemble the intent DNA fragment and the plasmid backbone for vector construction.

Results

Cloning eukaryotic genes by overlap PCR-based assembling of their exons

Firstly, we tried to clone the ORF of porcine *MSTN* gene (NC_010457) by the overlap PCR strategy. Myostatin is a myokine protein released by myocytes to inhibit myogenesis and the *MSTN* gene contains 3 exons and 3 introns (Fig.1A). Generally, to obtain the 1128 bp CDS, the exons were amplified respectively via PCR with porcine genomic DNA as template and were assembled one by one via overlap PCR. The principle for designing overlap PCR primers is illustrated in Fig.1B. The 3 exons were amplified with primer pairs of exon 1-F/exon 1-R, exon 2-F/exon 2-R, and exon

78 3-F/exon 3-R respectively (Fig.1B and 1D). The PCR products of these 3 exons were
79 purified by gel extraction, and further served as the templates for generating the whole
80 ORF fragment of *MSTN* gene (Fig.1C) by the second overlap PCR step
81 (Supplementary Table S2 and Fig.S2) with primers exon 1-F and exon 3-R (Fig.1B
82 and 1E). This technique makes gene cloning simply by exon-ligation through overlap
83 PCR without RNA extraction and RT-PCR, and can be applied widely for amplifying
84 various eukaryotic gene interrupted by introns for molecular biology study.

85 To demonstrate the feasibility, the chicken ovalbumin gene (NM205152, *OVA*)
86 containing 8 interval exons (with the codon ATG in exon 2) inlayed in the 7573 bp
87 genomic DNA was assembled similarly. Seven pairs of overlap PCR primers were
88 designed as described above. The 1161 bp ORF of *OVA* gene was generated by
89 assembling fragments from exon 2 to exon 8 (Supplementary Fig.S1A and S1B). We
90 next successfully assembled the 2132 bp ORF of human α -glucosidase gene with 11
91 exons (Supplementary Fig.S1C and S1D). Actually, in addition to the two-step
92 overlap PCR procedure described above, the intent ORF fragments of the two genes
93 could also obtained by one-step overlap PCR simply pooling all the overlap PCR
94 primers, the genomic DNA template and the PCR regents together in a single tube.
95 These studies suggested that the overlap PCR strategy is effective and practical for
96 amplifying long eukaryotic genes interrupted by introns, which may be more
97 convenient and economical when applied in the further studies.

98 **Cloning functional cassettes by overlap PCR-based assembling of different DNA**
99 **elements**

100 In order to further verify the advantage of the overlap PCR strategy, we tried to
101 assemble different DNA elements rapidly to generate designed functional cassettes.
102 Firstly, the SV40P.neo-IRES-tk-polyA cassette was designed and constructed. Three
103 pairs of overlapping primer were designed as shown in Fig.2A. The SV40P.neo, IRES
104 and tk-polyA fragments were amplified from plasmid DNA through touch-down PCR
105 [18] and purified by gel extraction routinely. The whole 3.3 kb fragment was
106 generated by overlap PCR with primers at the ends using the three fragments as the
107 templates (Fig.2B).

108 Secondly, we tried to assemble a larger Tet.SV40T-T2A-p53-polyA cassette,
109 which was intended to be 5.7 kb in length. Five pairs of overlapping primers were
110 designed as shown in Fig.2C. Corresponding DNA elements were PCR amplified
111 from either plasmid DNA or genomic DNA, and were assembled together
112 successfully by overlap PCR (Fig.2D). Furthermore, by using the overlap PCR
113 strategy, we also obtained the zinc finger cassette with three fingers for targeting the
114 *CCR5* gene (Fig.2E and 2F). The successful assembly of different DNA elements
115 through overlap PCR will facilitate the fast cloning of functional cassette by avoiding
116 multiple cloning steps.

117 **Fast plasmid vector construction by overlap-circled PCR**

118 In another attempt, we obtained a circle plasmid containing the selective marker gene
119 and the *MSTN* gene expression cassette through overlap-circled PCR (Fig.3). To
120 begin with, the backbone plasmid pBlueScript (Addgene) containing the *Amp^r* gene
121 and pUC replication initiation site was amplified with primers pBlue-F/pBlue-R. Then

122 the ORF of eukaryotic *MSTN* gene was cloned by the two-step overlap PCR strategy
123 as described above. Finally, the pBlueScript backbone fragment and the *MSTN* ORF
124 fragment were overlapped by circled PCR [14, 15] using the *pfu* DNA polymerase
125 without primers for generating the pBlueScript-MSTN construct, which was used to
126 transform the *E.coli* JM109 competent cells for the amplification. Then plasmid DNA
127 was extracted from culture of positive colonies, the digesting assay with *EcoR* I/*Xho* I
128 restrictive enzymes was conducted and agarose gel electrophoresis was performed to
129 confirm the positive plasmid clones. The results revealed that 4 out of the 12 clones
130 (33.3%) were positive. Here, we demonstrated a useful overlap-circled PCR method
131 for fast vector construction, through only two or three PCR procedures without using
132 any ligase and endonuclease.

133 **Discussion**

134 In summary, we firstly cloned the ORFs of genes porcine *MSTN*, chicken *OVA* and
135 human *α -glucosidase* by overlap PCR-based assembling their exons. Secondly, we
136 successfully obtained several functional cassettes by assembling different DNA
137 elements from plasmid or genomic DNA. Thus, the first advantage of the overlap
138 PCR strategy is that it is much convenient and cost-effective to conduct gene cloning
139 without RNA extraction and RT-PCR, and the cloning can be even achieved by one
140 step overlap PCR procedure. The second advantage is that functional cassettes can be
141 generated by quickly assembling corresponding DNA elements, facilitating the fast
142 cloning of complicated cassettes by avoiding multiple cloning steps. Finally, we
143 further demonstrated an interesting overlap-circled PCR method for fast vector

144 construction with only PCR procedures, which makes the vector construction much
145 easier and more economic by simply designing several pairs of primers without using
146 any ligase and endonuclease. To sum up, our work demonstrated that the overlap PCR
147 strategy can be flexible, effective, convenient and economical for gene cloning and
148 vector construction.

149 **Materials and Methods**

150 **Chemicals and stains**

151 *Taq* DNA polymerase, *Pfu* DNA polymerase, DNA marker, *E.coli* JM109 competent
152 cells were purchased from TransGen Biotech (Beijing, China). T4 DNA ligase and
153 restriction endonucleases were from NEB (Beijing, China).

154 **Preparation of genomic DNA**

155 The genome DNAs prepared from tissues (pig and chicken) was shown as follows and
156 that from cultured cells (human 293T) was manipulated similarly without the
157 digestion. Digest approximately 2 mm³ tissues with 300 µL digestion buffer [5 mM
158 EDTA, pH 8.0, 200 mM NaCl, 100 mM Tris, pH 8.0, 0.2% sodium dodecyl sulfate
159 (SDS)] with 0.4 mg proteinase K/1 mL digestion buffer, in a 1.5 mL tube at 55 °C
160 overnight. Add an equal volume (about 300 µL) of phenol/chloroform to each sample,
161 and vortex for 30 seconds. Centrifuge at room temperature at max speed for 5 minutes
162 to separate phases. Transfer the upper phase of each sample to a fresh tube. Add 1 mL
163 100% ethanol into each tube, mix completely. Centrifuge at max speed for 10 minutes,
164 and then pour out the ethanol. Wash the DNA pellets by adding 1 mL 70% ethanol
165 into the tubes, then centrifuge at max speed for 10 minutes. Pour out the ethanol from

166 the tubes, immediately dissolve the DNA pellets with 50 μ L TE buffer (10 mM
167 Tris-HCl, 0.2 mM Na₂EDTA, pH 7.5) after air-drying.

168 **Overlap PCR for assembling exons and DNA fragments**

169 The individual exons and DNA element fragments were amplified through standard
170 PCR procedure and reaction system (15 μ L) as shown in Supplementary Table2. The
171 overlap PCR for assembling exons and different DNA fragments were performed in
172 50 μ L reaction system (Supplementary Table2) using the touch-down PCR procedure
173 (Supplementary Fig.S2). Briefly, the touch-down PCR parameters used were as
174 follows: pre-denaturation at 95°C for 5 minutes; followed by 18 touch-down cycles
175 of denaturation at 95°C for 30 seconds, annealing at 68°C for 30 seconds with one
176 degree reduction in each cycle, and polymerization at 72°C for 30 seconds; then, 25
177 cycles of denaturation at 95°C for 30 seconds, annealing at 50°C for 30 seconds, and
178 polymerization at 72°C for 30 seconds; and a final incubation for 10 minutes at 72°C;
179 end with 10°C incubation. PCR products were purified by gel extraction using the kit
180 according to the manufacturer's instructions (OMEGA; Guangzhou, China). All
181 primers used in this work were ordered from Genscript (Nanjing, China) and the
182 sequence information was shown in Supplementary Table1.

183 **Vector construction by overlap-circled PCR**

184 For the plasmid vector construction by overlap-circled PCR, the plasmid backbone
185 and the insertion fragments with overlapping ends were firstly obtained by
186 respectively normal and overlap PCR amplification. Then, the two fragments were
187 overlapped by circled PCR [14, 15] using the *pfu* DNA polymerase without primers

188 for generating the intent construct. The product was directly used to transform *E.coli*
189 JM109 competent cells. Positive colonies were picked and cultured for plasmid DNA
190 extraction using the kit (OMEGA; Guangzhou, China). The digesting assay with
191 *EcoR* I/*Xho* I restrictive enzymes and agarose gel electrophoresis was conducted to
192 confirm the positive plasmid clones.

193 **Acknowledgements**

194 The authors would like to thank all the colleagues in Professor Zhang's lab for their
195 excellent technical assistance and helpful discussions.

196 **Competing interests**

197 No competing interests declared.

198 **Author contributions**

199 Zhongtian Liu and Tingting Zhang performed the experiments; Kun Xu designed the
200 experiments and wrote the manuscript; Tingting Zhang and Kun Xu funded this work.

201 **Funding**

202 This study was supported by grants from National Natural Science Foundation of
203 China (NSFC, 31501905) and Ph.D. Start-up Fund of Northwest A&F University
204 (2452015351).

205

247 **Figure legends**

248 **Figure1 Cloning the porcine *MSTN* gene by assembling of exons**

249 **A.** The schematic diagram of porcine *MSTN* gene in genome. The *MSTN* gene
250 contains three exons, which are illustrated by solid arrows in blue. The start codon in
251 exon 1 and stop codon in exon 3 labeled. **B.** The design of the overlapping primers.
252 Primers are boxed and their orientations are indicated by parallel arrows. Exon-exon
253 junctions are marked by vertical lines. The genome DNA sequence for the *MSTN* gene
254 is partially displayed. **C.** Schematic illustration of the *MSTN* cDNA generated by
255 assembling of exons. **D.** Agarose gel electrophoresis detection for three PCR
256 amplified exon fragments of *MSTN*. Lane M: DL200 marker, lane 1, 2, 3: exon 1,
257 exon 2 and exon 3. **E.** Agarose gel electrophoresis detection for the overlapped *MSTN*
258 cDNA fragment. Lane M: DL200 marker, lane 1: overlapped *MSTN*.

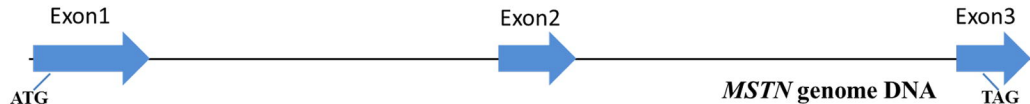
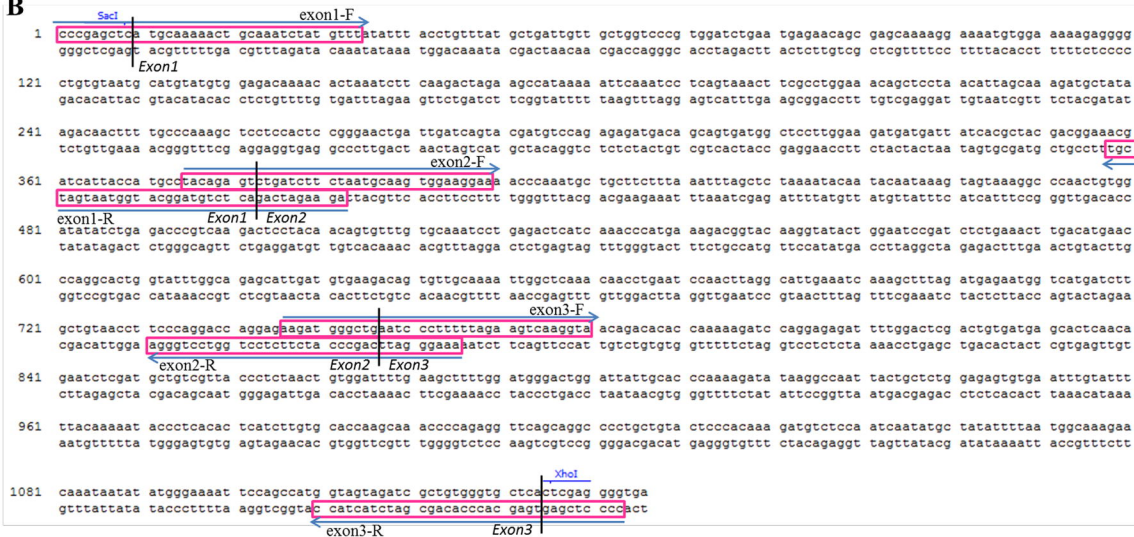
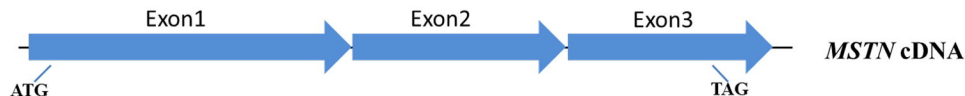
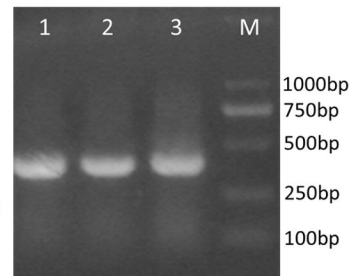
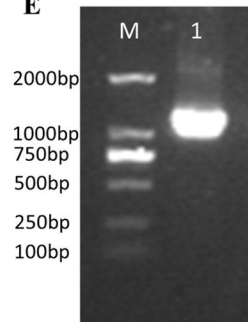
259 **Figure2 Cloning functional cassettes by assembling of different DNA elements**

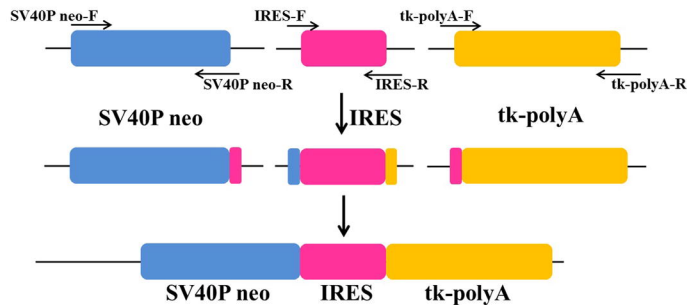
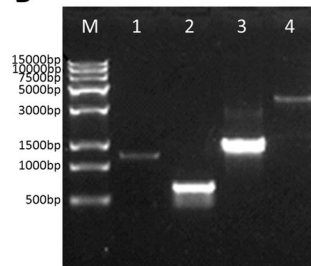
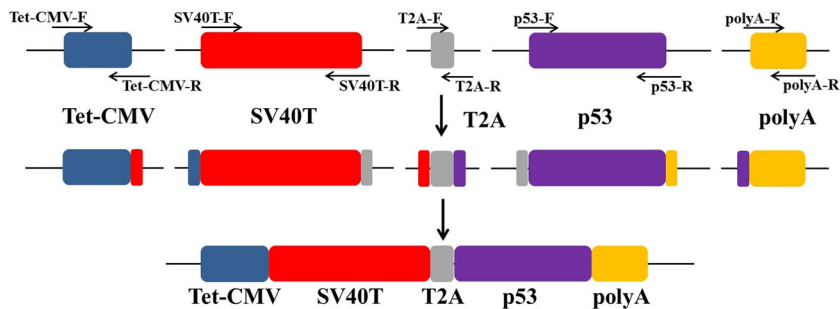
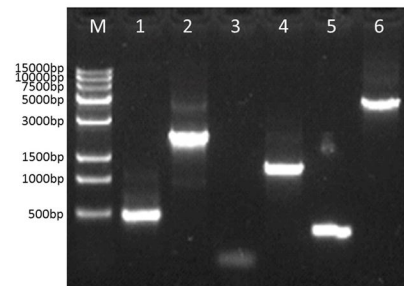
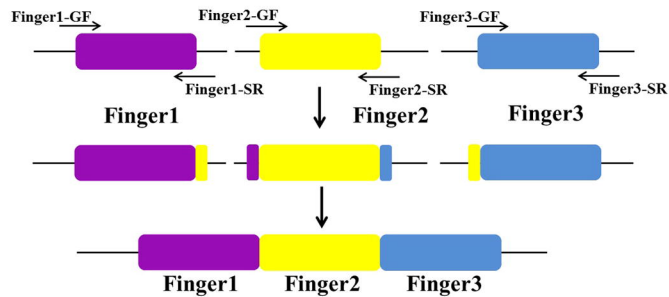
260 **A.** Schematic diagram for assembling the SV40P.neo-IRES-tk-polyA cassette. **B.**
261 Agarose gel electrophoresis detection of the PCR products. Lane M: Trans 15K
262 marker, lane 1, 2, 3, 4: represent respectively the PCR products of SV40P. neo, IRES,
263 tk-polyA and overlappedSV40P.neo-IRES-tk-polyA. **C.** Schematic diagram for
264 assembling the Tet.SV40T-T2A-p53-polyA cassette. **D.** Agarose gel electrophoresis
265 detection of the PCR products. Lane 1, 2, 3, 4, 5, 6: Tet.CMV, SV40T, T2A, p53,
266 polyA and overlapped Tet.SV40T-T2A-p53-polyA. **E.** Schematic diagram for
267 assembling the zinc finger cassette. **F.** Agarose gel electrophoresis detection of the
268 PCR products. Lane M: Trans 2K marker, lane 1, 2, 3, 4: Finger 1, Finger 2, Finger

269 and fused zinc fingers. All overlapping primers designed were labeled and indicated
270 by arrows.

271 **Figure3 Schematic illustration for the overlap-circled PCR-based plasmid vector**
272 **construction**

273 The *Amp^r* marker gene and the pUC replication initiation site of pBlueScript
274 backbone were amplified by normal PCR and the MSTN fragment was generated by
275 overlap PCR as described above. The pBlueScript backbone fragment and the MSTN
276 fragment were overlapped by circled PCR using the pfu DNA polymerase without
277 primers for generating the pBlueScript-MSTN construct. *E.coli* JM109 were
278 transformed with the product for cloning and the gel represents the result of digesting
279 assay with *EcoR* I/*Xho* I. Lane M: DL200 marker; lane 1-4: plasmids extracted from
280 positive JM109 transformants; the two intent bands should be 1168 bp and 1776 bp
281 respectively in length.

A**B****C****D****E**

A**B****C****D****E****F**