

1 **Rapid and scalable *in vitro* production of single-stranded DNA**

2 *Dionis Minev*^{1,2,3,4,#}, *Richard Guerra*^{2,3,4,#}, *Jocelyn Y. Kish*^{2,6,§}, *Cory Smith*^{2,5,§}, *Elisha Krieg*⁷,
3 *Khaled Said*^{2,5}, *Amanda Hornick*^{2,5}, *Hiroshi M. Sasaki*^{2,6}, *Gabriel Filsinger*^{2,5}, *Brian J.*
4 *Beliveau*^{2,6,†}, *Peng Yin*^{2,6}, *George M. Church*^{2,5}, and *William M. Shih*^{2,3,4,*}

5
6 ¹*John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge,*
7 *Massachusetts, USA;*

8 ²*Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston,*
9 *Massachusetts, USA;*

10 ³*Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA;*

11 ⁴*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,*
12 *Boston, Massachusetts, USA;*

13 ⁵*Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA;*

14 ⁶*Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, USA;*

15 ⁷*Leibniz Institute for Polymer Research Dresden e.V, Dresden, Germany;*

16 ^{#,§}*denotes equal contribution*

17 [†]*Current Address: Department of Genome Sciences, University of Washington, Seattle,*
18 *Washington, USA;*

19 ^{*}*Correspondence: William.Shih@wyss.harvard.edu*

20

21 **We present a rapid, scalable, user-friendly method for *in vitro* production of high-purity**
22 **single-stranded DNA (ssDNA) ranging from 89–3315 nucleotides in length. PCR with a**
23 **forward primer bearing a methanol-responsive polymer generates a tagged amplicon that**
24 **enables selective precipitation of the modified strand under denaturing conditions. We**
25 **demonstrate that the recovered ssDNA can be used for CRISPR/Cas9 homology-directed**
26 **repair in human cells, DNA-origami folding, and fluorescent in situ hybridization.**

27 DNA is instrumental to myriad applications in biological imaging, (bio)nanotechnology,
28 and synthetic biology. Many of the applications rely heavily on the availability of ssDNA.
29 Depending on the required size, scale, and purity, the production of ssDNA can become
30 prohibitively expensive or onerous. Although chemically synthesized ssDNA has become widely
31 available commercially, such DNA has an upper limit of ~200 nucleotides (nt)¹ in length and
32 often requires additional processing steps to remove impurities. Methods allowing the
33 production of ssDNA through a double-stranded DNA (dsDNA) template via enzymatic
34 processing², micro-bead sequestration³, rolling circle amplification⁴, or asymmetric PCR⁵ have
35 been introduced, but are often limited in either complexity of the protocols, scalability, and/or
36 purity of recovered strands. Recently, Palluk et al. have demonstrated a promising enzymatic
37 approach for de novo synthesis of ssDNA⁶. However, this method has yet to demonstrate
38 production of ssDNA >10 nt at high yields. Similarly, autonomous ssDNA synthesis via primer
39 exchange reaction (PER) is currently limited to lengths up to 60 nt⁷. Phagemid-based *in vivo*
40 production of ssDNA can yield biotech-scale quantities of arbitrary sequences, however the
41 method is less amenable to rapid prototyping due to increased lag time between sequence
42 design and strand production⁸. Thus a need persists for methods that allow for fast, user-
43 friendly, scalable, and low-cost *in vitro* production of ssDNA above 200 nt in length.

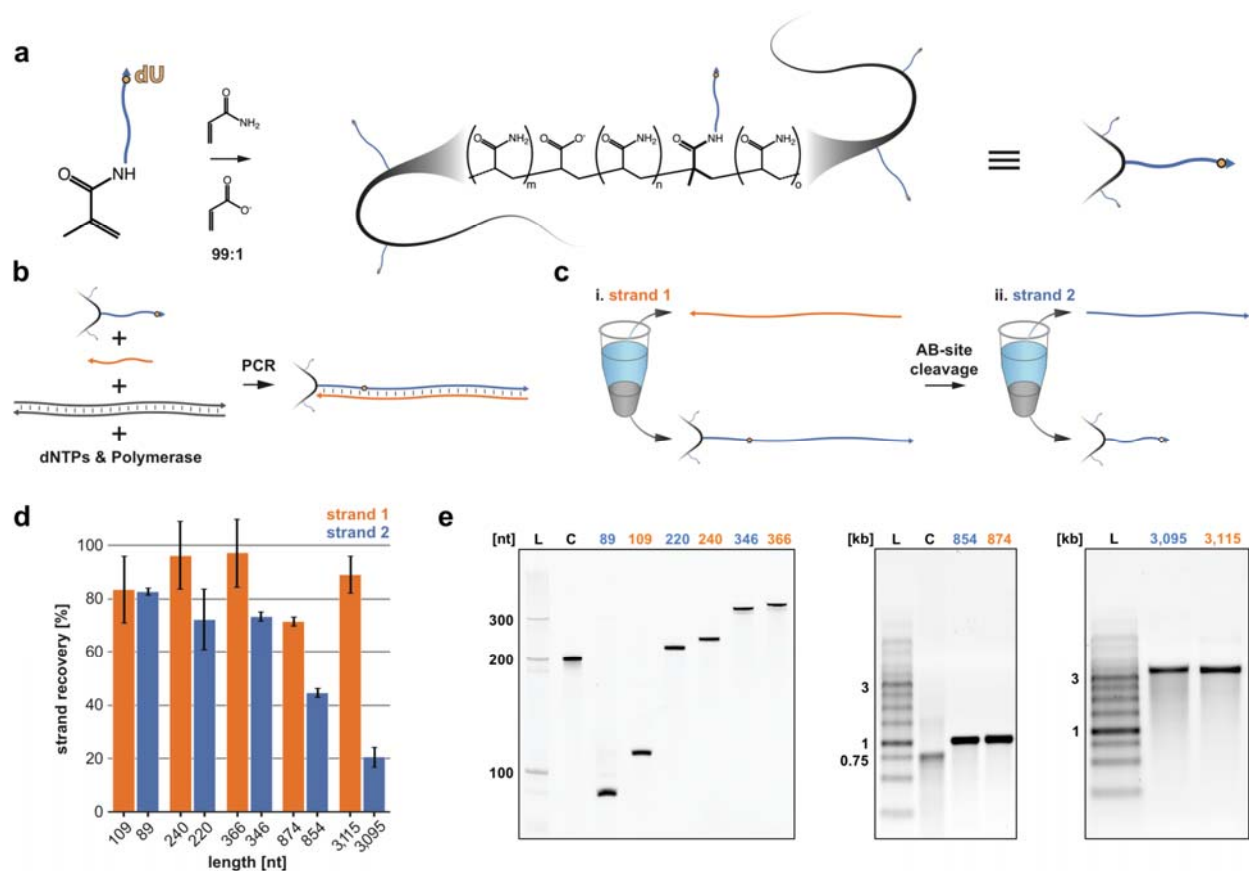
44 We present a method that we call Methanol-Responsive Polymer PCR (MeRPy-PCR),
45 inspired by our previous work of Krieg et al.⁹ (also see Krieg et al.¹⁰) We create a set of primers
46 bearing a linear polyacrylamide-co-acrylate tag by co-polymerizing a 5'-acrydite-modified primer
47 with acrylamide and sodium acrylate (Fig. 1a, Supplemental Fig. 1, Supplemental Tables 1, 2,
48 Supplementary Protocol 1). The modified primer can include a deoxyuridine (dU), which can be
49 placed anywhere along the sequence and allows the site-specific creation and subsequent
50 cleavage of an abasic site (AB site). We use the polymer-tagged primer in an otherwise
51 standard PCR reaction, resulting in a polymer-tagged amplicon (Fig. 1b, Supplementary
52 Protocol 2) that subsequently allows the selective precipitation (Supplementary Fig. 2) and

53 recovery of both sense and antisense strands away from each other (Fig. 1c, Supplementary
54 Fig. 3). Substitution of a polymer-tagged primer had no noticeably adverse effects on PCR
55 production in terms of strand yield and purity (Supplementary Fig. 4).

56 After PCR, we first recover untagged strand 1 in a supernatant by performing a
57 denaturing precipitation under alkaline conditions by addition of NaOH to 44 mM final
58 concentration, followed by mixing with one volume equivalent of methanol and then
59 centrifugation at 350–2,000 RCF (Fig. 1ci). We next recover complementary strand 2 by
60 resuspending the precipitated polymer-DNA pellet and incubating it with uracil-DNA glycosylase
61 (UDG) for 15 minutes to excise the dU nucleobase and create an AB site. The AB site is then
62 cleaved by incubating the polymer-DNA solution with 100 mM dimethylethylenediamine
63 (DMEDA)¹¹ for 15 minutes, followed by precipitation in 50% methanol to remove the waste
64 polymer-tagged DNA (Fig. 1cii). The procedure is completed within ~45–65 minutes (depending
65 on strand amplicon length), with strand 1 recovery accounting for the first ~15–25 minutes
66 (Supplementary Fig. 3).

67 We used this method to generate ssDNA ranging from 89–3115 nt in length by
68 amplifying an array of target sequences with MeRPy-PCR and recovering both strands 1 and 2
69 of each amplicon (Fig. 1d, Supplementary Figs. 5–9, Supplementary Note 1). The strand-
70 recovery protocol was nearly identical for all lengths and templates, apart from slight differences
71 in the alkaline denaturation step for the longest amplicons (see Supplementary Protocol 3).
72 Strand 1 was routinely recovered with a yield of 70% to >90% with respect to the initial MeRPy-
73 PCR amplicon. By contrast, recovery yield of strand 2 was lower as the length of the amplicons
74 increased (see Supplementary Note 2). We recorded absolute yields of ~2.2–12 pmol/100 μ L
75 PCR for strand 1 and ~0.5–12 pmol/100 μ L PCR for strand 2 (Supplementary Fig. 10,
76 Supplementary Yield Data). It should be noted that the final amount and purity of recovered
77 ssDNA depends on the efficiency and cleanliness of the PCR, therefore PCR optimization may

78 be desirable. Furthermore, we observed that ssDNAs recovered from MeRPy-PCR were of high
 79 purity, on par or better than a chemically-synthesized 200mer oligonucleotide after PAGE
 80 purification and an enzymatically produced 754mer oligonucleotide purchased from the
 81 commercial vendor Integrated DNA Technologies (Fig. 1e).



82
 83 *Figure 1: MeRPy-PCR overview and recovery yields for strands 1 (untagged) and 2 (initially*
 84 *tagged) of different amplicon lengths. (a) Production of the polymer-tagged primer. A 5'*
 85 *acrydite-modified primer is polymerized with acrylamide and sodium acrylate (ratio 99:1) to form*
 86 *a long linear DNA-tagged polymer. (b) MeRPy-PCR procedure following standard PCR*
 87 *guidelines. (c) i. Recovery of strand 1 under alkaline denaturing conditions and methanol*
 88 *precipitation. ii. Recovery of strand 2, after treatment with UDG and DMEDA followed by a*
 89 *methanol precipitation. (d) Recovery yield for strands 1 and 2 of various lengths. Bar graphs*
 90 *denoting the recovery yield (%). Strand recovery yield was determined by the absolute*

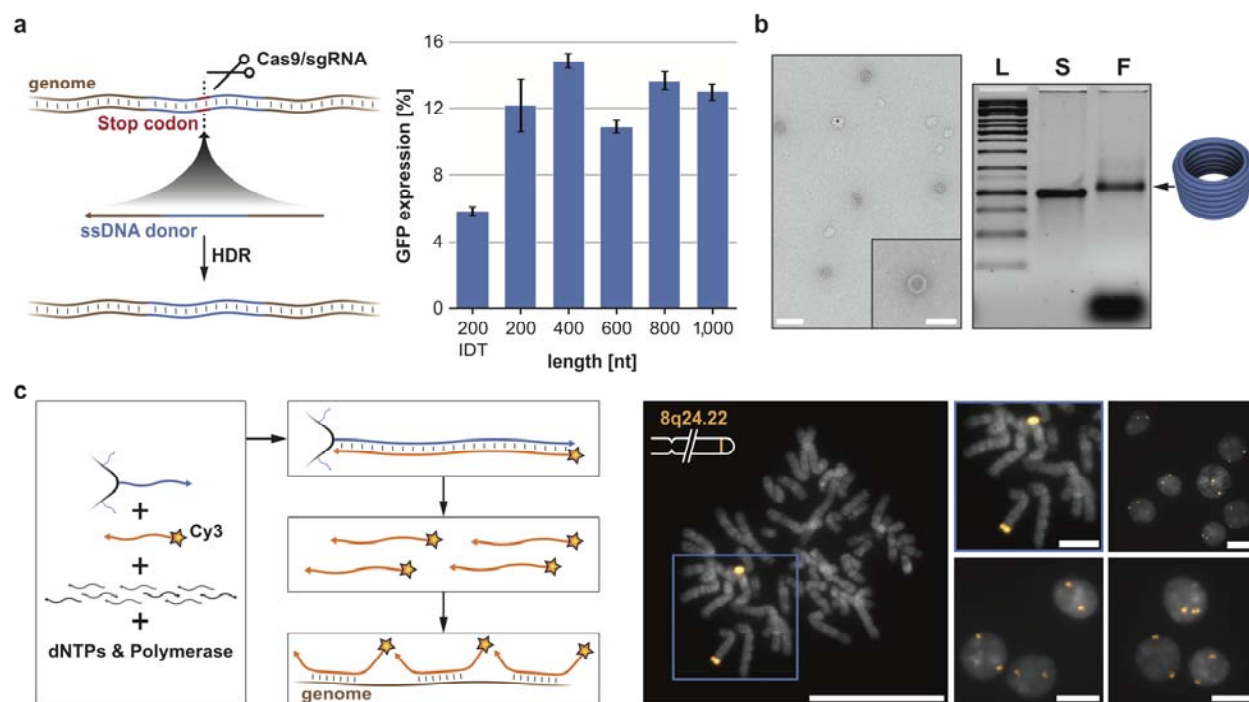
91 *recovered strand output (pmol) relative to MeRPy-PCR input (pmol). Data is shown as mean +/-*
92 *STD (N=3). (E) Gel electrophoresis of MeRPy-PCR derived ssDNA. Left, denaturing*
93 *polyacrylamide gel with L – 20 bp Ladder, C – 200mer control from Integrated DNA*
94 *Technologies (IDT). Middle and right, native agarose gels with L – 1 kb Ladder, C – 750mer*
95 *control from IDT. MeRPy-PCR derived and commercial ssDNAs were loaded with normalized*
96 *mass amounts for each gel lane in (E).*

97 To demonstrate the utility of MeRPy-PCR generated ssDNA for demand-meeting
98 applications, we show CRISPR/Cas9 mediated homology directed repair (HDR) in human cells,
99 fluorescent in situ hybridization (FISH) imaging, and DNA-origami folding. We picked the
100 untagged strand 1 for each application, based on the higher overall recovery yield and briefer
101 protocol. Each of the three tested applications utilizes ssDNA in varying capacities; DNA origami
102 requires long ssDNA scaffolds (>1 kb)¹²⁻¹⁴, FISH requires a library of >100 nt Cy3-labeled
103 strands to tile specific regions of the genome¹⁵, and CRISPR/Cas9 directed HDR has seen
104 growing interest in the field to use long ssDNA over dsDNA donors¹⁶⁻¹⁸, which can be difficult to
105 produce or else prohibitively expensive to purchase at sufficient scale for cell-culture
106 experiments.

107 For HDR, we assessed the performance of MeRPy-PCR generated ssDNA donors
108 (Supplementary Fig. 11) of varying size, relative to a purchased chemically synthesized 200 nt
109 donor from IDT. The ssDNA donor-mediated HDR removed a stop codon from a broken GFP
110 expression vector, restoring the GFP sequence and expression (Fig. 2a, Supplementary Fig. 12,
111 Supplementary Note 3). We generated 5 different ssDNA donors 200–1000 nt long, only varying
112 the homology-arm length. We produced the ssDNA donors at yields of ~13–34 pmol/100 μ L
113 PCR (Supplementary Table 3). The efficiency of HDR was comparable for the different MeRPy-
114 PCR generated ssDNA and the 200 nt chemically synthesized donor.

115 Next, we tested the ability to produce custom scaffolds for DNA-origami folding. DNA
116 origami is often limited to a defined number of ssDNA scaffolds based on the availability of
117 different M13 phage genomes. There is growing interest in the field for the design and
118 production of new scaffolds that offer a larger range of sequence space¹⁹. To address this
119 application, we first used MeRPy-PCR to generate a 3315 nt ssDNA derived from p7308 M13
120 genome (Supplementary Table 4). Using the produced ssDNA scaffold (1.35 pmol/100 μ L PCR,
121 Supplementary Table 5) we demonstrated folding of a DNA-origami barrel²⁰ (Fig. 2b,
122 Supplementary Figs. 13 and 14).

123 Finally, we demonstrated the ability to use MeRPy-PCR to generate a large library of
124 FISH probes with a Cy3 modified 5' end (Fig. 2c, Supplementary Fig. 15). We were able to
125 generate ~70 pmol/100 μ L PCR of Cy3-modified ~130 nt FISH probes (Supplementary Table 6),
126 that can successfully be used in imaging a distinct locus of the genome (Chromosome
127 8q24.22)¹⁴. As expected by the FISH probe design, we observed two puncta per cell, with the
128 puncta located towards the end of two similarly sized, medium-length chromosomes (Fig. 2c,
129 Supplementary Fig. 16). Use of MeRPy-PCR here highlights the ease with which FISH probes
130 can be generated in sufficient quantities for imaging, obviating the need for expensive and time-
131 consuming purifications.



132
 133 *Figure 2: Applications using ssDNA of various lengths. (a) Genome editing in human cells using*
 134 *CRISPR/Cas9. (left) A genomically integrated GFP coding sequence is disrupted by the*
 135 *insertion of a stop codon and a 68-bp genomic fragment from the AAVS1 locus. Restoration of*
 136 *the GFP sequence by HDR with a ssDNA donor sequence results in GFP+ cells that can be*
 137 *quantified by FACS. (right) Bar graph depicting HDR efficiencies induced by MeRPY-PCR*
 138 *derived ssDNAs of different lengths vs. a 200mer chemically synthesized strand from IDT. Data*
 139 *is shown as mean +/- STD (N=3). (b) ssDNA scaffold was generated via MeRPY-PCR from the*
 140 *phage genome, p7308, and used in the folding of a 30 nm DNA origami barrel. Agarose gel*
 141 *electrophoresis (right) shows the purified scaffold strand (S) alongside the folded barrel*
 142 *structure (F). Transmission electron microscopy depicts the folded origami (left). Scale bars*
 143 *denote (left) 100 nm and (right) 50 nm. (c) A library comprising 42,000 probe sequences*
 144 *designed to tile along an 8.4 Mbp region of Human Chromosome 8 was amplified from a small*
 145 *amount of template using MeRPY-PCR with a Cy3-labeled reverse primer and subsequent*
 146 *recovery of fluor-tagged strand 1 library. The generated fluor-labeled ssDNA library was*

147 *validated in situ on fixed human metaphase spreads and interphase cells. Scale bars denote 20*
148 *μm (zoom of metaphase spread scale bar denotes 5 μm).*

149 In summary, we have demonstrated that MeRPy-PCR can be performed without the need for
150 additional optimization beyond that needed for PCR in general, and can be used to recover high
151 yields of both forward and reverse strands, with a briefer protocol and higher yields for the
152 untagged strand. We further demonstrated that the generated ssDNA can be used in a variety
153 of demand-meeting applications in synthetic biology, (bio)nanotechnology, and biological
154 imaging. The short time frame to recover the strands is user-friendly and lowers the bar to rapid
155 in-house production of large quantities of ssDNA. Importantly, the low-cost production of strands
156 via MeRPy-PCR may enable the accelerated exploration of scaffold design space in DNA
157 origami, of genome visualization with FISH, and of the efficiency and off-target effects of single
158 stranded donor DNA in CRISPR/Cas9 HDR.

159

160 References:

- 161 1. Beaucage, S.L., & Caruthers, M.H. *Tetrahedron Lett.* **20**, 1859-1862 (1981).
- 162 2. Kujau, M.J., & Wöfl, S. *Mol. Biotechnol.* **7**, 333-335 (1997).
- 163 3. Bowman, B.H., & Palumbi, S.R. *Methods Enzymol.* **224**, 399-406 (1993).
- 164 4. Schmidt, T.L., *et al. Nat. Comm.* **6**, 8634 (2015).
- 165 5. Veneziano, R., *et al. Sci. Rep.* **8**, 6548 (2018).
- 166 6. Palluk, S., *et al. Nat. Biotechnol.* **36**, 645-650 (2018).
- 167 7. Kishi, J.Y., *et al. Nat. Chem.* **10**, 155-164 (2017).
- 168 8. Praetorius, F. *et al. Nature* **552**, 84-87 (2017).
- 169 9. Krieg, E. & Shih, W.M. *Angew. Chem. Int. Ed. Engl.* **57**, 714-718 (2018).
- 170 10. Manuscript in preparation, to be submitted to BioRxiv shortly
- 171 11. McHugh, P.J., & Knowland, J. *Nucleic Acids Res.* **23**, 1664-1670 (1995).

- 172 12. Rothemund, P.W. *Nature* **440**, 297-302 (2006).
173 13. Douglas, S.M., *et al.* *Nature* **459**, 414-418 (2009).
174 14. Douglas, S.M., *et al.* *Nucleic Acids Res.* **37**, 5001-5006 (2009).
175 15. Beliveau, B.J., *et al.* *Proc. Natl. Acad. Sci. USA* **109**, 21301–21306 (2012).
176 16. Richardson, C.D., *et al.* *Nat. Biotechnol.* **34**, 339-344 (2016).
177 17. Davis, L., & Maizels, N. *Cell Rep.* **17**, 1872-1881 (2016).
178 18. Roth, T.L., *et al.* *Nature* **559**, 405-409 (2018).
179 19. Chen, X., *et al.* *ACS Appl. Mater. Interfaces.* **10**, 24344-24348 (2018).
180 20. Ponnuswamy, *et al.* *Nat. Comm.* **8**, 15654 (2017).

181

182 **Acknowledgments:** The authors would like to thank Christopher M. Wintersinger for fruitful
183 discussions and for providing a 3D rendered figure of the DNA-Origami barrel. This work was
184 funded by support from ONR Award N00014-18-1-2566 and the Wyss Institute at Harvard Core
185 Faculty Award. Research reported in this publication was supported by National Human
186 Genome Research Institute of the National Institutes of Health under award number
187 RM1HG008525, the Uehara Memorial Foundation Postdoctoral Fellowship, a Damon Runyon
188 Cancer Research Foundation Postdoctoral Fellowship (HHMI).

189

190 **Author Contributions:** Single-stranded DNA production via MeRPy-PCR: D.M., R.G., and
191 W.M.S. designed the experiments. D.M. and R.G. performed the experiments. D.M. and R.G.
192 analyzed the data. DNA origami folding and imaging: D.M., R.G. and W.M.S. designed the
193 experiments. D.M. and R.G. performed the experiments. CRISPR/Cas9 HDR: C.S. designed the
194 experiments. C.S., K.S., and A.H. performed the experiments. C.S., and K.S. analyzed the data.
195 FISH: J.Y.K., H.M.S. and B.J.B. designed the experiments. H.S.M. and B.J.B. designed the
196 sequence library. H.M.S. performed the emulsion PCR. J.Y.K. performed the FISH experiments.
197 D.M., R.G., J.Y.K., C.S. and W.M.S. wrote the paper.

198

199 **Competing Interests:** A provisional US patent has been filed based on this work. G.M.C. is a

200 co-founder of Editas Medicine and has other financial interests listed at

201 arep.med.harvard.edu/gmc/tech.html. P.Y. is co-founder of Ultivue Inc. and NuProbe Global.