SUPPORTING INFORMATION

A zinc finger protein array for the visual detection of specific DNA sequences

for diagnostic applications

Moon-Soo Kim¹, Gulnaz Stybayeva², Ji Youn Lee², Alexander Revzin², David J. Segal^{1,†}

¹Department of Pharmacology and Genome Center, ²Department of Biomedical Engineering,

451 Health Sciences Drive, University of California, Davis, CA 95616

Sequences of hairpin DNA target oligonucleotides

All sequences written 5' to 3' with ZFP binding site shown in bold.

Target site1 for rrsA27 and rrsA62: GAC GAT TGA ACG CTG GCG GCA GGC CTA ACA CAT GCA AGT GGG TTTT CCC ACT TGC ATG TGT TAG GCC TGC CGC CAG CGT TCA ATC GTC

Target site 2 for rrsA125 and rrsA160: GAC TGG GAA ACT GCC TGA TGG AGG GGG ATA ACT ACT GGA GGG TTTT CCC TCC AGT AGT TAT CCC CCT CCA TCA GGC AGT TTC CCA GTC

Target site 3 for rrsA1175 and rrsA1192: GAC GAG GAA GGT GGG GAT GAC GGG TTTT CCC GTC ATC CCC ACC TTC CTC GTC

Hairpin oligonucleotide of Zif268/PBSII site, used as an irrelevant control: GGC TTT CCA CAC CGC CCA CGC GGG TTTT CCC **GCG TGG GCG GTG TGG AAA** GCC

AAATTGAAGA	GTTTGATCAT	GGCTCA <mark>GATT</mark>	GAACGCTGGC	GGCAGGCCTA
ACACATGCAA	GT CGAACGGT	AACAGGAAGA	AGCTTGCTTC	TTTGCTGACG
AGTGGCGGAC	GGGTGAGTAA	TGTC TGGGAA	ACTGCCTGAT	GGAGGGGGAT
AACTACTGGA	AACGGTAGCT	AATACCGCAT	AACGTCGCAA	GACCAAAGAG
GGGTACCTTC	GGGCCTCTTG	CCATCGGATG	TGCCCAGATG	GGATTAGCTA
GTAGGTGGGG	TAACGGCTCA	CCTAGGCGAC	GATCCCTAGC	TGGTCTGAGA
GGATGACCAG	CCACACTGGA	ACTGAGACAC	GGTCCAGACT	CCTACGGGAG
GCAGCAGTGG	GGAATATTGC	ACAATGGGCG	CAAGCCTGAT	GCAGCCATGC
CGCGTGTATG	AAGAAGGCCT	TCGGGTTGTA	AAGTACTTTC	AGCGGGGAGG
AAGGGAGTAA	AGTTAATACC	TTTGCTCATT	GACGTTACCC	GCAGAAGAAG
CACCGGCTAA	CTCCGTGCCA	GCAGCCGCGG	TAATACGGAG	GGTGCAAGCG
TTAATCGGAA	TTACTGGGCG	TAAAGCGCAC	GCAGGCGGTT	TGTTAAGTCA
GATGTGAAAT	CCCCGGGCTC	AACCTGGGAA	CTGCATCTGA	TACTGGCAAG
CTTGAGTCTC	GTAGAGGGGG	GTAGAATTCC	AGGTGTAGCG	GTGAAATGCG
TAGAGATCTG	GAGGAATACC	GGTGGCGAAG	GCGGCCCCT	GGACGAAGAC
TGACGCTCAG	GTGCGAAAGC	GTGGGGAGCA	AACAGGATTA	GATACCCTGG
TAGTCCACGC	CGTAAACGAT	GTCGACTTGG	AGGTTGTGCC	CTTGAGGCGT
GGCTTCCGGA	GCTAACGCGT	TAAGTCGACC	GCCTGGGGAG	TACGGCCGCA
AGGTTAAAAC	TCAAATGAAT	TGACGGGGGC	CCGCACAAGC	GGTGGAGCAT
GTGGTTTAAT	TCGATGCAAC	GCGAAGAACC	TTACCTGGTC	TTGACATCCA
CGGAAGTTTT	CAGAGATGAG	AATGTGCCTT	CGGGAACCGT	GAGACAGGTG
CTGCATGGCT	GTCGTCAGCT	CGTGTTGTGA	AATGTTGGGT	TAAGTCCCGC
AACGAGCGCA	ACCCTTATCC	TTTGTTGCCA	GCGGTCCGGC	CGGGAACTCA
AAGGAGACTG	CCAGTGATAA	ACTG <mark>GAGGAA</mark>	GGTGGGGATG	AC GTCAAGTC
ATCATGGCCC	TTACGACCAG	GGCTACACAC	GTGCTACAAT	GGCGCATACA
AAGAGAAGCG	ACCTCGCGAG	AGCAAGCGGA	CCTCATAAAG	TGCGTCGTAG
TCCGGATTGG	AGTCTGCAAC	TCGACTCCAT	GAAGTCGGAA	TCGCTAGTAA
TCGTGGATCA	GAATGCCACG	GTGAATACGT	TCCCGGGCCT	TGTACACACC
GCCCGTCACA	CCATGGGAGT	GGGTTGCAAA	AGAAGTAGGT	AGCTTAACCT
TCGGGAGGGC	GCTTACCACT	TTGTGATTCA	TGACTGGGGT	GAAGTCGTAA
CAAGGTAACC	GTAGGGGAAC	CTGCGGTTGG	ATCACCTCCT	ТА
	AAATTGAAGA ACTACTGCAA AGTGGCGGAC GGTACCTTC GTAGGTGGGG GGATGACCAG GGATGACCAG GCAGCAGTGG CGCGTGTATG AAGGGAGTAA TTAATCGGAA TTAATCGGAA GATGTGAAAT CTTGAGTCTC TAGAGATCTG TGACGCTCAG GGCTTCCGGA AGGTTAAAAC GTGGTTTAAT CGGAAGTTTT CTGCATGGCT AACGAGCGCA AAGGAGACTG ATCATGGCCC AAGAGAAGCG TCCGGATTGG TCGTGGATCA GCCCGTCACA	AAATTGAAGAGTTTGATCATACACATGCAAGTAGTGGCGGACGGGTGAGTAAAACTACTGGAAACGGTAGCTGGGTACCTCGGGCCTCTGGGTAGGTGGGCTAACGGCTACGGATGACAGGGAATATTGCGGAGCAGTGGGAGAAGGCTAAGGGAGTAAAGTTAATACCCACCGGCTAATTACTGGGCGTAAATCGGAACCCCGGGCTGCACCGGCTAACTCCGTGCCATTAATCGGAATTACTGGGCGGATGTGAAATCCCCGGGCTGTGAGAGATCTCGTAGAGGGGGTAGACGCTCAGGTGCGAAAGCTGACGCTCAGACCCAGGAAGCAGGTTAAAACTCAAAGAGCGCGGGAAGTTTCAGAGACGACGGGAAGCTGGTCGTCAGGACGGGAAGCCTTACGACGAGAAACGAGACGAACCCTTACCAGAAGGAAAGCGACCTCGCAAGAAGAGAAGCGACCTCGCAAGAAGAGAAGCGACCTCGCAAGAAGGAGACGACCTCGCAAGAAGGAAAGCGACCTCGCAAGACGGGATTAGAATGCCACGAAGGAGACGACCTCGCAAGACGGGAACCACAGAATGCCACGACGGGAAGCACACGAATGCCACGACGGGAGCCGTACGGAGAC	AAATTGAAGAGTTTGATCATGGCTCAGATTACACATGCAAGTAACAGGAAGAAGTGGCGGACGGGTGAGTAATGTCGGGTACCTCCGGGCTACTTGCCATCGGATGGGGTACCTCCGGGCCTCTTGCCATGGCACAGGATGACCAGCCACACTGGAACTGAGACACGGAGCAGTGGGGAATATGCACCAGGCGAGCAGCAGTGGGGAATATGCACCAGGCGACACCGGCTAAAGTTAATACCTTTGCTCATTCACCGGCTAACTCCGTGCCAGCAGCCGCGGTTAATCGGAATTACTGGGCGTAAAGCGCACCACAGGATAACCCCGGCCCAACCTGGGAACTTGAGATCTCGTAGAGAGGGGTAGAATTCCTAATCGGAACTCCGGGCCCAACCTGGGAACTGAGATCTGGAGGAATACCGTGGGCAAGATGACGCTCAGGTGCGAAGACGTGGGAGACACTGACGCTCAGGCTAAACGACTTAAGTCGACCGGCTTCCGGAGCCAAACGACGCGAAGACCCGGAAGTTTCCAGTGAAATGACGGGGCGGCTTACAGCGCGAAGAACCCGGAAGACCCGGAAGACCCCAGTGAAAACTGGGAGAACTGGCAAGGCACCCTTACCAGGCTACACACAAGGAAGCGACCTCGCAAACCGGGAGAAAACAGAGACGACCTGCGAAACCACACACACAAGAAAGCGACCTGCGAAACCACACACACAAGAAAGCGACCTGCGAAAACCACACACACAAGAAAGCGACCTGCGAAAACCACACACACAAGAAAGCGACCTGCAAAACGAACCACACAAGAAAGCGACCTGCGAAAACCAACCACACAAGAAAGCGACCTGCGAAACCAACACCACACAAGAAAGCGACCTGCGAAACCAACACCACACAAGAAAGCG<	AAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCACACATGCAAGT <gaacggt< td="">AACAGGAAGAAGCTTGCTTGAGTGGCGGACGGGTGAGTAATGTCTGGGAAACCGCCGGAAGGGTACCTTCGGGCCTCTTGCCATCGGAACGGTCCAGATGGTAGGTGGGGTAACGGCTCACCTAGGCACGGTCCAGACGGGATGACCAGCCAACATGGAACTGCAGACGGGACCAGTGGCAGCAGTGGGGAATATTGCACAGGCTGACGGCCTCAGATCGCGTGTATGAAGAAGCCTTCGGGTTGTAAAGTACTTCAAGGAGATAAAGTTAATACCTTTGCTCATTGACGTTACCCACCGGCTAACTCCGTGCCAGCAGCCGCGTAATACGGAGTAATCGGAATTACTGGGCCAACAGGCACAGCGGCCCCTGAGGTAAATCCGTGGGAAACGCGGCCCCTGCAGAGATACGAGGTCAAGCTCCGGGCTCAACAGGAATACGCGGCCCCCTTGAAGATCCGTGGGAAAGCGCGGCCCCCTGTGAGAAGCTGACGCTAGGTGGAAAGCGCGGCCCCCTGGGGTTGCAGAGGGTTAAAACCATGCGAAGCGGAGAGACTAACTGGCGGGGTTAAAATCAAAGCAAGCGGAGACCTAACTGGCCGGGAAGTTATCGAGAGACTTGACTGGGCGCGGAACGGGGGAAGTTATCGAGAGACCCCGGAAGACGCGGGCCCGCAAGGAGACGGCCTATACCTTGTGTGCAGGGTGGGAAGGGGTACACCGCGGTACACACGGGTGGGAAGGCTCACAAAGGCGAGGCAACCTCGCGAAACCTGGGGAACGCGGGCACGGGCAAGGCCAACCTGCGAAAACCGGGGCACGGGTGCAAAAGGGTAGACCAGCCTGCGAAAACCGGGGAAAGCCGGGCACGGGAAGTTACC</gaacggt<>

Figure S-1. Location of three target regions (27-62 bp, 125-160 bp, and 1175-1192 bp) in *rrsA* gene (1542 bp). The target regions are high-lighted in yellow, light green, and cyan for ZFP rrsA27 and rrsA62, ZFP rrsA125 and rrsA160, and ZFP rrsA1175 and rrsA1192, respectively.



Figure S-2. (A) ZFP microarray to detect the immobilized GZF1 on the slide by Cy5-labeled secondary antibody. ZFP GZF1 fused with MBP (Maltose binding protein) was printed on the PEG gel slide. The slide was blocked with ZBA/1% BSA (Right panel) and washed with ZBA and air-dried. The slide was incubated with primary antibody anti MBP-HRP and washed with ZBA and ZBA/0.05% TWEEN. Incubation with Cy5-labeled mouse IgG secondary antibody and washing were performed as in the case of the primary antibody. The slide was scanned with an Agilent Microarray Fluorescent Scanner. H6-HS2 is also ZFP fused with MBP. (B) ZFP microarray for detection of GZF1 binding to target DNA labeled with Cy5. ZFP GZF1 was printed on the PEG gel slide which was then incubated with Cy5-labeled with target DNA. After washing with ZBA and air-drying, the slide was scanned an Agilent Microarray Fluorescent Scanner. Another ZFP, H6-HS2, designed to bind a different target site was used as a negative control. GZF1 concentration was 1.9 μM and its lower concentration was 1.3 μM.



Figure S-3. Spotted ZFP microarrays using the SEER-LAC system detect *E. coli* DNA sequences in a concentration-dependant manner. MBP-LacA-ZFP rrsA125 was printed on the PEG gel slide using a manual arraying tool. Each array (square) contains 20 spots (5 x 4 rows). The arrays were then incubated for 20 min with MBP-ZFP rrsA160-LacB and target DNA oligonucleotides (1: DNA 2.5 μ M, 2: DNA 0.25 μ M, 3: DNA 5 nM, and 4: No DNA). The slide was washed with ZBA and air-dried, followed by incubation with nitrocefin.



Figure S-4. Electromobility shift assay (EMSA) of engineered zinc finger proteins. Exemplary data are shown for (A) LacA-rrsA27, (B) LacA-rrsA125, and (C) rrsA160-LacB. The concentrations of the protein are indicated. Arrows, bound probe; asterisks, free probe.