Supplementary Online Data for

A genotype-to-phenotype map of *in vitro* selected RNA-cleaving DNAzymes: Implications for accessing the target phenotype

Kenny Schlosser[1], Jeffrey C.F. Lam[1] and Yingfu Li[1,2*]

Department of Biochemistry and Biomedical Sciences[1] and Department of Chemistry[2], McMaster University, Hamilton, Ontario, Canada L8N 3Z5

* Corresponding Author:

1200 Main Street West

Hamilton, Ontario

Canada L8N 3Z5

Tel: 905-525-9140 ext. 22462

Fax: 905-522-9033

Email: liying@mcmaster.ca

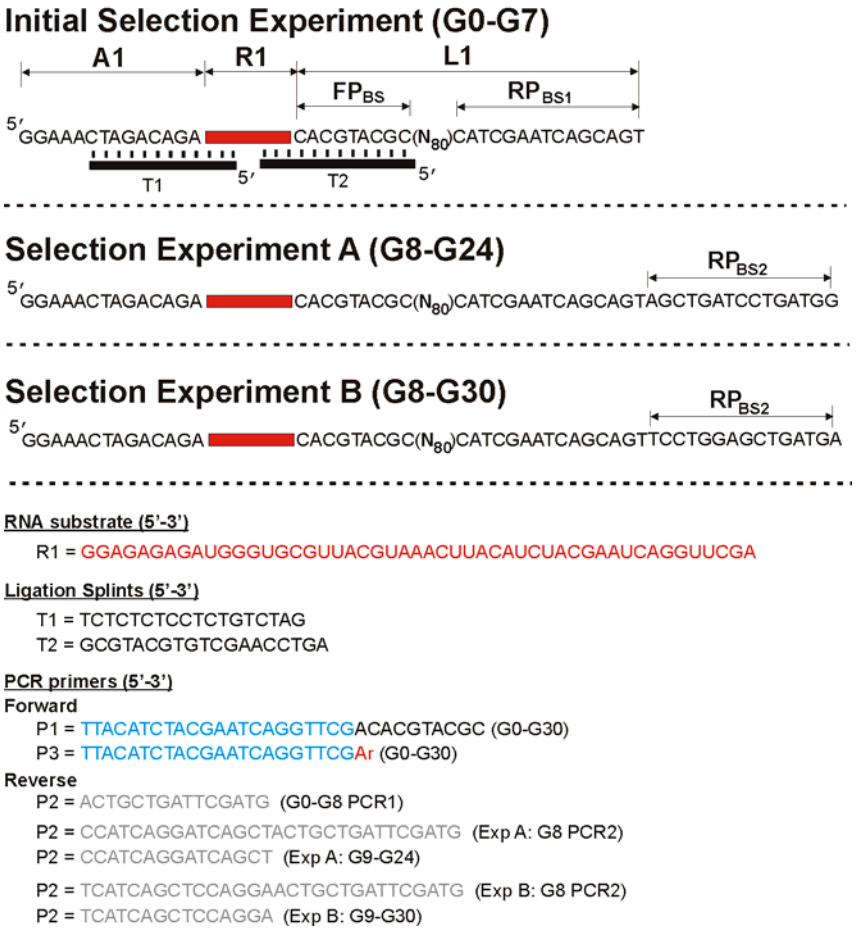

Figure S1 Library Design and Relevant Sequences. Each molecule in the initial library contains three key domains: a 14-nt DNA fragment (denoted A1) precedes a 50-nt RNA fragment (denoted R1), which in turn precedes another 104-nt DNA fragment (denoted L1). L1 contains 80 random-sequence nucleotides (N_{80}, where N = A, G, C or T) that serve as the putative catalytic domain and provides the initial sequence diversity. The N_{80} region is flanked by a 9-nt forward primer binding site (FP_{BS}) and a 15-nt reverse primer binding site (RP_{BS1}), which are used for PCR amplification. Different reverse primer binding sites (RP_{BS2}) and corresponding P2 primers, were used in experiment A and B to minimize cross-contamination of the amplifying DNA species. The A1, R1, and L1 fragments are ligated by T4 DNA ligase in the presence of ligation splints T1 and T2. Sequences denoted by P1, P2, and P3 serve as primers during PCR (Ar = adenine ribonucleotide).

E18-G15-C1A	CACGTACGCATTTCCAGCGGATCGATTCTCTTTCCCGTCGTAGGTATGACCAGGGAAGAATAGGTGGACATAAATTGATGGTGTCGGGCATCGAATCAGTAGTAGCTGATCCTGATGG
E18-G15-C6A	CACGTACGCACTTCCAGCGGATCGATTCTCTTTCCCGTCGTGGGTATGACCAGGGAAGAATGGGTGGACACAAATTGATGGTGTCGGGCGTCGAATCAACAGTAGCTGATCCTGATGG
E18-G21A	CACGTACGCACTTCCAGCGGATCGATTCTCTTTCCCGTCGCGGGTATGACCAGGGAGGAATGGGTGGACACAGATTGATGGTGTTGGGCATCGAATTGGTGGTAGCTGATCCTGATGG
E88-G24A	CACGTACGCACTTCCAGCGGATCGAAATCTCTCTTTGCAGCTGGACTCGGAGGCCTGCTTCCACCAGTAGGGGGGTTTGTTCAGGGTGGTATCGAATCGGTAGTAGCTGATCCTGATGG
E25-G15A	CACGTACGCTACTCTCAGTGAGGCGAAATCTTCTCTCTGCGGGAACAATCGGGGGCGCAGTGATCAAGGGTGGAAATGGGGATGGGTGCATCGAATTAGCAGTAGCTGATCCTGATGG
E25-G24A	CACGTACGCACTGTCAGCGGACTCGAAATCTTCTCTCTGCGGGAACAGTTGGGGGCGCAGTGATCAAGGGTGGAAATGGGGATGGGTGCATCGAATTGGCAGTAGCTGATCCTGATGG
E75-G15A	CACGTACGCACTGCCAGCGGCGCGAGGCTCTTGATCGGGCGCAGGAGGGGACCGGTGATATCGGCATCCTCGATGTTAGACTGGATGGTCATCGAATTAGCAGTAGCTGATCCTGATGG
E75-G21A	CACGTACGCACTGTCAGTGACGCGAGGCTCTTGATCGGGCGCAGGAGGGGACCGGTGATACCGGCATCCTTGATGTTAGGCTGGATGGGTATCGAATTGGCATAGCTGATCCTGATGG
E78-G18A	CACGTACGCACTTCCAGCGGATCGATTCTCACCCTGTGGCGTATTATCGTTGGATGATGGTGGTGGGCGGAGGCTTTCGTACATTGGACATTGAATCGGCAGTAGCTGATCCTGATGG
E42-G15-C16A	CACGTACGCACTTCCAGCGGATCGAAATCTTGAACGCAGCTAGGTCTCGGGTGTGGCGGTGAGTTGGCGTAGGCCATGCCTTCCGCTGGCATCGAATCATCAGTAGCTGATCCTGATGG
E42-G15-C24A	CACGTACGCACTTCCAGCGGATCGAAATCTCGAACGCAGTTAGGTCTCGGGTGTGGCGGTGAGTTGGCGTAGGCCATGTCTTCCGCTGGCATCGAATTAGCAGTAGCTGATCCTGATGG
E14-G10A	CACGTACGCACTGCTAGCAGCCCGAAATCGCTCTCCCAATATGGTCTTCCGGGGAAGACGGTAATAGGAGAAATGGTGCCTTGTCCGTGCATCGAATCAGTAGTAGCTGATCCTGATGG
E14-G15A	CACGTACGCACTGCTAGCAGCTCGAAATCGCTCTCTCAATATGGGCTTTCGGGGAAGACGGTAATAGGAGAAATGGTGCCTTGTCCGTGCATTGAATCAGCAGTAGCTGATCCTGATGG
E14-G24A	CACGTACGCACTGCTAGCAGCTCGAAATCGCTCTCCCAATGTGGCCTTCCGGGGAAGACGGTAATAGGAGAAATGGTGCCCTGTCCGTTCATCGAATCCGCCCCAGCTGATCCTGATGG
E74-G15-C30A	CACGTACGCACCTCAATAGCAGCGTTAACAAAAAGTTTCGAGAAAGCGAATCATCTAACGGTGGTGACTCCATTGGTTTTTTGGGTGGGCATCGAATTAGCAGTAGCTGATCCTGATGG
E74-G15-C33A	CACGTACGCACCTCAATAGCAGCGCTAACAAAAAGTTTCGAGAAAGCGAATCATCTAACAGTGGTGACTCCATTGGTCTTTTGGGTGGGCATCGAATTAGCAGTAGCTGATCCTGATGG
E76-G15A	CACGTACGCACTGCCAGTGGCGCGAATTCTCTGGGAGATCTGTATAGGGTTGCCTGCGAGTTGACAGGGATGGTGTGCAGTTTGTGTGGCATCGAATCAACAGTAGCTGATCCTGATGG
E23-G15A	CACGTACGCACTGCCAGCGGCGCGATTCACTGTCGGAGACTAGTTGTTGCCTTCGGCTTGGAAGGACAAAACTTTTGTCATAGCGTGGGCATTGAATCAGTAGTAGCTGATCCTGATGG
E77-G13A	CACGTACGCACTGCCAGCGGCGCGATTCTGTTTCGGCGTGGTTATACAGCTTGAGTGGTGGCACTCTTGCCAGCCTAAGTGTTGGGGTGCATCGAATTAGTAGTAGCTGATCCTGATGG
E105-G10A	CACGTACGCACTGCCAGCGGCGCGATTTGGCTCTACCGTTCATTCGGCTGTGTTGCCCTAGGTAGGGTTCAGCAGTGATTTTGGTGGGGCATCGAATTAGCAGTAGCTGATCCTGATGG
E2-G8A	CACGTACGCACTGTGAGCCGACACGAATCTCTGAGTCTTCGGTCGAGAGCCTGCTATTCAGCATTAGGAGGCGCGCTCTCGGCTGGTGGCATCGAATTAGCAGTAGCTGATCCTGATGG
E129-G7	CACGTACGCATTTCCAGCGGATCGAATTGTCATCTCATGAGCGTAGGGAGGTCTATTGTTGGGGGGCAGTGTATGATAGCAGGTGAGCATCGAATCAGCAGT
E126-G7	CACGTACGCACTGCCAGCGGCGCGAACTTACCCTCATTGGTGCAGTAATAAGAGCAGATGTTGCGTTGTTCGGGAACGTAATGACGGTGGCATCGAATCAGCAGT
E121-G7	CACGTACGCATTACGAGCCGTACGAGTCTCTCTTAAGGTGTGTTGCTCGTATGTCATCTAGCTGTGCGCGGGGGGGTGGTCGAGTGTGGCATCGAATCAGCAGT

Figure S2. Sequences of specific DNAzyme clones isolated from experiment A. The sequence class (E) and generation (G) in which each clone was found is indicated. In cases where more than one clone was assayed from the same generation and sequence class, a further numerical designation is provided to distinguish between them (C). The PCR amplification efficiency was determined for clones highlighted in yellow.

E53-G30B	CACGTACGCATTGCCAGCGGCGCGAATCGACTCTGGGCGGATCTGGTCACCGGACCAAGGACGGAGGATTTGGGAGGTCTAGTTTAGGTATCGGATCGGTGGTCCTGGAGCTGATGA
E25-G10B	CACGTACGCTACTCTCAGTGAGGCGAAATCTTCTCCCTGCGGGAACTATCGGGGGCGCAGTGATCAGGGGTGGATATGGGGATGGGTGCATCGTATCATGTAGTTCCTGGAGCTGATGA
E25-G30B	CACGTACGCACTCCCAGTGGGGCGAAATCTTCTCTCTGCGGGAACAGTTGGGGGCGCGGTGATCAAGGGTGGAAATGGGGATGGGTGCATCGAGTTGGTGGTCCTGGAGCTGATGA
E52-G30B	CACGTACGCCAGGCAGGGTCGAGGTGGGATCGGATGATGTATTGTAACGGCACTCTCAGTGAGGCGGGTCGCTCTCCGTGGTTGGGTATCGAGTCAGTGGTCCTGGAGCTGATGA
E58-G30B	CACGTACGCATTGTCAGCAGGCAGGCGAGGGACCATGGGTGGGGCACTCGGGCCGGGACTCGAATCGTCTCGCCTGTAGGTCGGTTTGCCATTGAATCGGGGTCCTGGAGCTGATGA
E54-G27B	CACGTACGCGGAGGGGACGGAGGGTGGGGGGCGTATTGCACTCTTAGCGAGCCGAAATCTCGGGAGGCTGGGTTTGTGTTTTGTGGGCATCGAATCCGTGGTCCTGGAGCTGATGA
E60-G30B	CACGTACGCACTCCCAGCGGGACGAGTCTCAGTCAGGAGGCAGCCGGGAGTAGGTTGGAAGGGCTCTGAACACCTGGTTGGGTTGGTTGCATCGAGTCATGGTCCTGGAGCTGATGA
E56-G30B	CACGTACGCTGGATCGGACCGGAGAGGGAGGATGTGAGTGCGTAACCCACTGCCAGCGGCGCGATTGTCACCCCTCGTCTAGAGTTGGCTATCGAATCATGGTCCTGGAGCTGATGA
E8/9-G27B	CACGTACGCACAGCCAGTGGCTCGAGTCTGCTCTTCGCTGGGATAGGGGCTACGAGGCGGGGGCTACTGACGAGTAGGTAGCTTGGGGGCATCGATCGCGGTCCTGGAGCTGATGA
E51-G30B	CACGTACGCTGGAGGAGACGGCGGCTTGGGGGTGTGGACCACCGGGCCTACGATGGGGTTATCTAGGATGGATAGGGTTTGGATGGCATCGAATCGGGGTCCTGGAGCTGATGA
E63-G30B	CACGTACGCATTGTCAGCGACACGATTCTTGGCGGGATGTCTACGTTACGCTGGTCAGGAGCTGGAGCGCAGGTTCGGAGGTGGGCTGGCATCGAGTTGGTGGTCCTGGAGCTGATGA
E9-G21B	CACGTACGCACTGCCAGCGGCACGATTCTTGGGTGTGATGTGATATGGGCTACGAGGCGGGGGCTACTGAAGAGTAGGTAGCTTGGGGGTATTGAATCGCGGTCCTGGTGCTGATGA
E8-G13B	CACGTACGCACAGCCAGCGGCTCGAGTCTGCTCCTCGCTGGGATAGGGGTTCATGTTGAAACACTAACCTAAATTTCGACGGGTTGTGGCATCGAATCGCAGTTCCTGGAGCTGATGA
E8-G21B	CACGTACGCACAGCCAGTGGCTCGAGTCTGCTCTTCGCTGGGATAGGGGCTCACGTTGAAACACTAACCTGAGTTTCGACGGGTTGTGGCATTGATCAGTGGTCCTGGAGCTGATGA
E2-G18B	CACGTACGCACTGTGAGCCGACACGAATCTCTGAGTCTTCGGTCGAGAGCCTGCTATCCAGCATTAGGAGGCGCGCTCTCGGCTGGTGGCATTGAATCATGGTCCTGGAGCTGATGA
E2-G21B	CACGTACGCATTGTGAGCCGACTCGAATCTCTGAGTCTATGGTCGAGAGCTTGCTGTCCAGCATTAGGAGGTGCGCTCTCGGCTGGTGGCATTGAATCGGGGTCCTGGAGCTGATGA
E5-G24B	CACGTACGCACTGTTAGTAACACGAGTCACCTCCTCGGCAGAGCTGCAGTCATCGGCTTCGTGAGGGGGTGTAGACCTGTTGGGGGGGCATCGAATCGGGGTCCTGGAGCTGATGA
E3-G15B	CACGTACGCACTCCTAGCAGGGCGAAATCTTTCTGGCGAGAGCTGAACCTTAAGTCCTGGACTTAGAGGTTCTGTGACTGTGGCCGGCATTGAATCAGTAGTCCTGGAGCTGATGA
E38-G18B	CACGTACGCAATCTTAGCAAGACGAGTCTCCTCTTCTGGCGTTGGGCACATGTTGGAGAACACTGCCCAGTTTTGGTGAATACATGTGGCATCGAATCATGGTCCTGGAGCTGATGA
E17-G10B	CACGTACGCACTAATAGCATTACGAATCCGACACAAGACGTTGGAAAATTTGGAAGTCTGAGGGGCTTTTGTTTGTAGTCGTCCTTGGCATCGAATCAGTAGTTCCTGGAGCTGATGA
E4-G13B	CACGTACGCACTATGAGCCATACGACGTCCCTCTCCAGCGCCGCATGAAGCAGAACGTAGGTGAAGGATCCCTTAAGTGGCTCGGGTGGCATCGAATAGCAGTTCCTGGAGCTGATGA
E4-G18B	CACGTACGCACTATGAGCCATACGATGTCTCTCTCCAGCGCCGCATGTAGCAGTTACGTAGGTGTAGGATCCTTTTAGTGGCTCGGGTGGCATTGTAATCATGGTCCTGGAGCTGATGA
E36-G15B	CACGTACGCACTTCCAGCGGATCGAAATGTCCTATCTCCTTGGCCGTTGGAGCCGGGAGGGACAATAGCAGATTTGTATGATGTCGGTGCATCGAATCAGTAGTCCTGGAGCTGATGA
E1-G15B	CACGTACGCACTCCCAGTGGGGCGAAATTAACTTGAATGATGGTGCTGTTGGCTCAGCTTACGCACACGAGAGGGTTTGGTGGTTGGGCATCGAATAGCAGTCCTGGAGCTGATGA
E14-G10B	CACGTACGCACTGCTAGCAGCCCGAAATTGCTCTCCCAATATGGGCTTCCGGGGAAGACGGTAATAGGAGAAATGGTGCCTTGTTTGTGTATCGAATCAGTAGTCCTGGAGCTGATGA
E42-G15B	CACGTACGCACTTCCAGCGGATCGAAATCTCGAACGCAGTTAGGTCTTGGGTGTGGCGATGAGTTGGCGTAGGCCATGCCTTCCGCTGGTATTGAATCAGTAGTCCTGGAGCTGATGA
E43-G15B	CACGTACGCACTATCAGCGATACGAAGACTATGCTCCTATGCTATGCCTGACATTAATGGGGTTAGTCCCCATGTTGGGCGGGTTGGCCATCGAATTAGCAGTCCTGGAGCTGATGA
E18-G15B	CACGTACGCATTTCCAGCGGATCGATTCTCTTTCCCGTCGCAGGTATGACCAGGGAAGAATGGGTGGACACAAATTGATGGTGTTGGGTATCGAATCAGTAGTCCTGGAGCTGATGA
E28-G10B	CACGTACGCATTGTCAGCGACACGAGTCTACGCTCTTTGATCAGACGGATATGCAGTGCTTGGAAGGTGAAATCTTCTGACGTTGTGCCATACGAATTAGCAGTCCTGGAGCTGATGA
E23-G13B	CACGTACGCACTGCCAGCGGCGCGATTCACTGTCGGAGACTAGTTGTTGCCTTCGGCTTAGAAGGACAAAACTTTTGTCATAGCGTGGGCATCGAATAGCAGTTCCTGGAGCTGATGA
E30-G13B	CACGTACGCACTGTCAGCGACACGAAATCTTCCTGGTAAAAGGTCTCGCCTTCTGAGTTTAGGTCAATGTGACCGTTTCTCGTTGGGGGCATCGATCAGTAGTTCCTGGAGCTGATGA
E173-G13B	CACGTACGCACTTTCAGCGAATCGAATTATATCTCGAGTGAATAGTACTCCGCCATGGGAACTGGGGTGGGGTGAACAGAAGCTTTGGGCATCGATCAGTAGTCCTGGAGCTGATGA
E46-G10B	CACGTACGCACTATCAGCGATGCGAATTCTCATACAGAATGTGTTTACTCCGAGGTAGGTGTATGGATTGTAGATCGCTAGTTCTGTGGCATCGAATCAGTAGTCCTGGAGCTGATGA
E198-G8B	CACGTACGCACACTCAGCGAGGCGAGTCTCTCTACCAGCCTGGTAGGAAGTTATAGGAGGCGATTATAAGTCGCGGGGAGTGGCCATGGCATCGAATCAGTAGTTCCTGGAGCTGATGA

Figure S3. Sequences of specific DNAzyme clones isolated from experiment B. The specific sequence class (E) and generation (G) in which each clone was found is indicated. The PCR amplification efficiency was determined for clones highlighted in yellow.