Journal of Visualized Experiments

Exploring sequence space to identify binding sites for regulatory RNA-binding proteins

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59635R2
Full Title:	Exploring sequence space to identify binding sites for regulatory RNA-binding proteins
Keywords:	RNA-binding protein; iterative selection amplification; splicing; polypyrimidine-tract/3'splice site; SELEX; sequence evolution in a test tube
Corresponding Author:	Ravinder Singh University of Colorado Boulder Boulder, CO UNITED STATES
Corresponding Author's Institution:	University of Colorado Boulder
Corresponding Author E-Mail:	ravinder.singh@colorado.edu
Order of Authors:	Ravinder Singh
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Boulder, CO, USA

TITLE:

Exploring Sequence Space to Identify Binding Sites for Regulatory RNA-Binding Proteins

AUTHORS & AFFILIATIONS:

Ravinder Singh[1]

[1]Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, CO 80309

Corresponding author:

Ravinder Singh (rsingh@colorado.edu)

KEYWORDS:

RNA-binding protein, iterative selection amplification, splicing, polypyrimidine-tract/3'splice site, SELEX, sequence evolution in a test tube

SUMMARY:

Sequence specificity is critical for gene regulation. Regulatory proteins that recognize specific sequences are important for gene regulation. Defining functional binding sites for such proteins is a challenging biological problem. An iterative approach for identification of a binding site for an RNA-binding protein is described here and is applicable to all RNA-binding proteins.

ABSTRACT:

Gene regulation plays an important role in all cells. Transcriptional, post-transcriptional (or RNA processing), translational, and post-translational steps are used to regulate specific genes. Sequence-specific nucleic acid-binding proteins target specific sequences to control spatial or temporal gene expression. The binding sites in nucleic acids are typically characterized by mutational analysis. However, numerous proteins of interest have no known binding site for such characterization. Here we describe an approach to identify previously unknown binding sites for RNA-binding proteins. It involves iterative selection and amplification of sequences starting with a randomized sequence pool. Following several rounds of these steps—transcription, binding, and amplification—the enriched sequences are sequenced to identify a preferred binding site(s). Success of this approach is monitored using in vitro binding assays. Subsequently, in vitro and in vivo functional assays can be used to assess the biological relevance of the selected sequences. This approach allows identification and characterization of a previously unknown binding site(s) for any RNA-binding protein for which an assay to separate protein-bound and unbound RNAs exists.

INTRODUCTION

In cell biology, gene regulation plays a central role. At one or multiple steps along the gene expression pathway, genes have the potential to be regulated. These steps include transcription (initiation, elongation, and termination) as well as splicing, polyadenylation or 3' end formation, RNA export, mRNA translation, and decay/localization of primary transcripts. At these steps, nucleic acid-binding proteins modulate gene regulation. Identification of binding sites for such

proteins is an important aspect of studying gene control. Mutational analysis and phylogenetic sequence comparison have been used to discover regulatory sequences or protein-binding sites in nucleic acids, such as promoters, splice sites, polyadenylation elements, and translational signals[1–4].

Pre-mRNA splicing is an integral step during gene expression and regulation. The majority of mammalian genes, including those in humans, have introns. A large fraction of these transcripts is alternatively spliced, producing multiple mRNA and protein isoforms from the same gene or primary transcript. These isoforms have cell-specific and developmental roles in cell biology. The 5' splice site, the branch-point, and the polypyrimidine-tract/3' splice site are critical splicing signals that are subject to regulation. In negative regulation, an otherwise strong splice site is repressed, whereas in positive regulation an otherwise weak splice site is activated. A combination of these events produces a plethora of functionally distinct isoforms. RNA-binding proteins play key roles in these alternative splicing events.

Numerous proteins are known whose binding site(s) or RNA targets remain to be identified[5, 6]. Linking regulatory proteins to their downstream biological targets or sequences is often a complex process. For such proteins, identification of their target RNA or binding site is an important step in defining their biological functions. Once a binding site is identified, it can be further characterized using standard molecular and biochemical analyses.

The approach described here has two advantages. First, it can identify a previously unknown binding site for a protein of interest. Second, an added advantage of this approach is that it simultaneously allows saturation mutagenesis, which would otherwise be labor intensive to obtain comparable information about sequence requirements within the binding site. Thus, it offers a quicker, easier, and less costly tool to identify protein binding sites in RNA. Originally, this approach (SELEX or Systematic Evolution of Ligands by EXponential enrichment) was used to characterize the binding site for the bacteriophage T4 DNA polymerase (gene 43 protein), which overlaps with the ribosome binding site in its own mRNA. The binding site contains an 8-base loop sequence, representing 65,536 randomized variants for analysis[7]. Second, the approach was also independently used to show that specific binding sites or aptamers for different dyes can be selected from a pool of approximately 10^{13} sequences[8]. In fact, this approach has been broadly used in many different contexts to identify aptamers (RNA or DNA sequences) for binding numerous ligands, such as proteins, small molecules, and cells, and for catalysis[9]. As an example, an aptamer can discriminate between two xanthine derivatives, caffeine and theophylline, which differ by the presence of one methyl group in caffeine[10]. We have extensively used this approach (SELEX or iterative selection-amplification) to study how RNA-binding proteins function in splicing or splicing regulation[11], which will be the basis for the discussion below.

The random library: We used a random library of 31 nucleotides. The length consideration for the random library was loosely based on the idea that the general splicing factor U2AF^{65} binds to a sequence between the branch-point sequence and the 3' splice site. On average, the spacing between these splicing signals in metazoans is in the range of 20 to 40 nucleotides. Another protein Sex-lethal was known to bind to a poorly characterized regulatory sequence near the 3'

splice site of its target pre-mRNA, *transformer*. Thus, we chose a random region of 31 nucleotides, flanked by primer binding sites with restriction enzyme sites to allow for PCR amplification and attachment of the T7 RNA polymerase promoter for in vitro transcription. The theoretical library size or complexity was 4^{31} or approximately 10^{18}. We used a small fraction of this library to prepare our random RNA pool (~10^{12}–10^{15}) for the experiments described below.

PROTOCOL:

NOTE: **Figure 1** provides a summary of key steps in the iterative selection-amplification (SELEX) process.

1. Generation of a random library template

1.1) Synthesize the forward primer 5'- GTAATACGACTCACTATAGGG*TGATCA*GATTCTGATCCA-3' and the reverse primer 5'- GCGAC*GGATCCAAGCTT*CA-3' by chemical synthesis on a DNA synthesizer.

NOTE: The primers and the random library can be synthesized commercially.

1.2) Synthesize a random library oligonucleotide template 5'-GGTGATCAGATTCTGATCCA(**N_1…N_{31}**)TGAAGCTTGGATCCGTCGC-3' by chemical synthesis. Use an equimolar mixture of four phosphoramidites during synthesis for the 31 randomized positions shown above as N.

NOTE: The sequence of the library template contains 31 random nucleotides (N1 to N31) and flanking sequences for the binding of the forward and reverse primers. The forward primer includes the T7 RNA polymerase promoter sequence (underlined) for in vitro transcription and a restriction site *Bcl1* (italicized) for cloning. The reverse primer contains restriction enzyme sites *BamH1* and *HindIII* (italicized) to facilitate cloning.

2. Generation of the DNA random library pool

2.1) Attach the T7 RNA polymerase promoter to the library by polymerase chain reaction (PCR) containing 1 µM DNA random library pool, 1 µM of each primer, 20 mM Tris (pH 8.0), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 µg/µL acetylated bovine serum albumin, 2 units of *Taq* polymerase, and 200 µM each of dNTPs (deoxyguanosine, deoxyadenosine, deoxycytidine, and deoxythymidine triphosphate).

2.2) Use five cycles of denaturation, annealing, and extension steps (94 °C for 1 min, 53 °C for 1 min, and 72 °C for 1 min) of PCR followed by one cycle of extension (72 °C for 10 min).

3. Synthesis of pool 0 RNA

3.1) Set up a 100 µL transcription reaction[12]. Mix T7 transcription buffer, 1 µM random library

pool DNA, 10 mM dithiothreitol (DTT), 2 mM guanosine triphosphate (GTP), 1 mM each of adenosine triphosphate (ATP), cytidine triphosphate (CTP), and uridine triphosphate (UTP), and 2 units/µL T7 RNA polymerase.

NOTE: RNA can be transcribed in vitro using commercially available kits with an option of the T7 or SP6 RNA polymerase.

3.2) Incubate the above reaction mixture in a microcentrifuge tube for 2 h at 37 °C.

3.3) Gel purify RNA in a 10% denaturing polyacrylamide gel.

3.4) Identify location of the transcripts on the gel by staining it with methylene blue or autoradiography by including traces of radioactivity (0.5 µL or less of α-^{32}P UTP) in the transcription reaction.

3.4.1) Place the gel slice in a centrifuge tube and break into smaller pieces, for example, with a homogenizer tip. Add proteinase K (PK) buffer (100 mM Tris, pH 7.5, 150 mM NaCl, 12.5 mM EDTA, 1% Sodium dodecyl sulfate) to immerse the gel pieces. Leave the tube on a nutator from 2 h to overnight at room temperature.

3.5) Spin in a high-speed microcentrifuge (14,000 rpm or 16,873 x *g*) for 5 min at room temperature to remove the gel debris and recover the buffer solution.

3.6) Vortex the sample two times with an equal volume of phenol-chloroform and one time with chloroform.

3.7) Mix the aqueous phase from above with one-tenth volume of sodium acetate (3.0 M, pH 5.2), 10 µg of tRNA or 20 µg of glycogen, and ethanol (2–3 volumes, stored at -20 °C). Leave the tubes at -80 °C for 1 h.

3.8) Spin the tubes containing the solution for 5–10 min at 4 °C in a microcentrifuge (14,000 rpm or 16,873 x *g*). Discard the supernatant carefully. Rinse the RNA pellet with 70% ethanol and spin for 2–5 min. Aspirate ethanol carefully. Air dry the RNA pellet.

3.9) Solubilize the RNA pellet in 50 µL of water treated with of diethyl pyrocarbonate (DEPC). Leave the sample at -20 °C for storage.

NOTE: Purify the RNA using commercially available spin columns, which are currently more commonly used to remove unincorporated radioactivity and serve as a quick and more convenient alternative for RNA purification.

CAUTION: Use an acrylic glass shield, gloves, and other precautions to protect from radioactivity.

4. Protein binding reaction and separation of bound RNA

4.1) Carry out binding of protein and RNA in 10 mM Tris-HCl, pH 7.5 in a volume of 100 µL by adding the following ingredients to these final concentrations: 50 mM KCl, 1 mM DTT, 0.09 µg/µL bovine serum albumin, 0.5 units/µL RNasin, 0.15 µg/µL tRNA, 1 mM EDTA, and 30 µL of appropriate recombinant protein (PTB) concentration. Add RNA from the appropriate pool.

NOTE: The splicing factor U2AF^{65} typically binds to the polypyrimidine-tract/3' splice sites of model introns with a binding affinity (equilibrium dissociation constant or K_d) of approximately 1–10 nM. Therefore, the first two rounds of binding used protein concentration 10-fold above the K_d for U2AF^{65}; for SXL and PTB proteins, the starting concentration in this range was only our best guess. This ensured that desired RNA species that could bind, although lower affinity sequences also potentially bound. In rounds 3 and 4 (transcription, binding, and amplification), the protein concentration was reduced three-fold (Step 4.1). This was done to successively eliminate low affinity RNA species.

4.2) Place the tubes containing the binding reactions for about 30 min at 25 °C in a temperature block (or on ice).

4.3) Fractionate the bound RNA from the unbound RNA for the first 4 rounds of selection-amplification using the following steps.

4.3.1) Filter the sample (100 µL) at room temperature through a nitrocellulose filter attached to a vacuum manifold.

NOTE: The RNA-protein complex, but not the unbound RNA, remains on the filter.

4.3.2) Chop the filter with retained RNA into fragments with a sterile razor blade; insert these into a centrifuge tube. Recover RNA by tumbling the tube gently for a minimum of 3 h (or overnight) with filter pieces immersed in the proteinase K (PK) buffer.

4.3.3) Deproteinize the RNA sample by vortexing it in the presence of an equal volume of phenol-chloroform (1:1) and then of chloroform. Recover the aqueous phase each time by centrifuging the sample at high speed for 5 min at room temperature.

4.3.4) Mix it with sodium acetate (0.1 volume of 3.0 M, pH 5.2) and ethanol (2–3 volumes of absolute ethanol, 200 proof). Leave the tube in a -80 °C freezer for 30 min, centrifuge it at high speed for 10 min, and, following the washing and drying steps (step 3.8), solubilize the RNA in water treated with DEPC. These steps are outlined above (steps 3.6 to 3.9).

4.4) Separate the protein-bound RNA fractions from the unbound fractions for the last 2 rounds (rounds 5 and 6; transcription, binding, and amplification) as follows. Reduce the protein concentration in the binding reaction (step 4.1) by further three-fold for additional selection pressure to enrich high-affinity binding sequences and preferentially remove low-affinity sequences.

4.4.1) Pre-cast a native polyacrylamide gel (5% with 60:1 acrylamide:bis-acrylamide ratio) in 0.5x TBE buffer (Tris-Borate-EDTA) prior to setting up the above RNA:protein-binding (step 4.1) reaction. Electrophorese this gel in a cold room (4 °C) by applying 250 V for 15 min.

4.4.2) Pipette the above RNA:protein binding reactions (step 4.1) into different wells of this gel.

NOTE: The protein is stored at -80 °C and diluted prior to use in 20 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), pH 8.0, 20% glycerol, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 0.05% NP-40, and 1 mM dithiothreitol (DTT). Addition of 0.5–1.0 mM protease inhibitor phenylmethane sulfonyl fluoride (PMSF) is optional. In the binding reaction, this buffer contributes about 6% glycerol, which allows direct loading of samples into the wells without the need for mixing them with a separate gel-loading buffer.

4.4.3) Fractionate the bound RNA from the unbound RNA using gel electrophoresis in a cold room (4 °C) at 250 V for 1 to 2 h. This process is also known as gel mobility shift assay.

NOTE: Duration of electrophoresis varies depending on features of the RNA and protein used for binding.

4.4.4) Expose the gel to an X-ray film and identify the location of the bound RNA using autoradiography. Cut out the gel slice with the bound RNA and insert it into a tube.

4.4.5) Incubate the crushed gel slice in the PK buffer used for elution for 3 h or overnight.

4.4.6) Repeat steps 3.5 to 3.9 outlined above. Briefly, vortex the eluted RNA sample vigorously first with phenol-chloroform and then with chloroform.

4.4.7) Mix sodium acetate and ethanol with the aqueous phase after chloroform extraction. Following incubation in -80 °C freezer, spin at 4 °C for 5–10 min to collect the RNA pellet. Wash the RNA pellet with ethanol and air dry it by leaving the lid of the tube open. Dissolve the RNA in water treated with DEPC.

NOTE: Switching to the gel mobility shift assay for fractionation allows elimination of unwanted RNA species that might have been enriched for binding, for example, to the nitrocellulose filter here (or any matrix) used in the initial rounds for fractionation.

5. Reverse transcription and PCR amplification

5.1) Synthesize cDNA from the dissolved RNA using reverse transcriptase and the reverse primer by incubating the 20 µL reaction (2 µL of 10x RT Buffer, 2 µL of AMV reverse transcriptase, 1 µM reverse primer, 10 µL of RNA, RNase inhibitor optional) at 42 °C for 60 min.

5.2) Amplify the cDNA using 20–25 PCR cycles as described in step 2.1.

6. Transcription and protein binding

6.1) Repeat the process of RNA synthesis, protein binding, separation of protein bound and unbound fraction, as described in section 3–5 above.

7. Analysis of RNA-protein interactions

7.1) Use the gel mobility shift assay (step 4.4) or the filter binding assay (step 4.3) to determine binding affinity and specificity for selected pools or individual sequences within each pool (see step 8.3).

7.2) Use autoradiography or a phosphor imager to detect and quantify bands in the bound fraction and unbound fraction.

8. Cloning and sequencing

8.1) Digest the final PCR DNA product with restriction enzymes Bcl1 and HindIII for 1–2 h, ligate with the appropriately digested pGEM3 or other plasmids carrying the restriction sites from 2 h to overnight, transform the ligation product into competent bacterial cells by heat shock or electroporation using standard molecular biology procedures[13].

8.2) Grow bacteria overnight by plating the transformed cells on agar plates with Luria-Bertani (LB) medium and ampicillin (50 µg/mL) at 37 °C. Pick colonies to inoculate culture tubes containing LB liquid medium with ampicillin and grow at 37 °C in a shaking incubator overnight. Purify plasmid DNAs containing DNA inserts using a standard plasmid isolation protocol[13].

NOTE: Commercial kits are available for plasmid purification.

8.3) Sequence the plasmids with DNA inserts using the dideoxy chain termination sequencing protocol, following manufacturer's instructions for sequencing[14].

NOTE: Sequencing can be performed in house or done commercially.

9. Sequence alignment

9.1) Align sequences and obtain a consensus binding site(s) using available online alignment tools (https://www.ebi.ac.uk/Tools/msa/).

REPRESENTATIVE RESULTS:

The following observations demonstrate successful selection-amplification (SELEX). First, we analyzed pool 0 and the selected sequences for binding to the protein used for the iterative selection-amplification approach. **Figure 2** shows that the mammalian polypyrimidine-tract binding protein (PTB) shows barely detectable binding to the pool 0 sequence but high affinity

for the selected sequence pool. There was barely detectable binding to pool 0 when we used about 300-fold higher protein concentration for binding than used for the selected pool. Thus, there was at least a several hundred-fold difference in protein binding affinities between the random or starting pool and the selected pool. This observation experimentally confirms that the selection-amplification protocol described here is successful.

Second, we sequenced the selected pool and determined a consensus binding site. The consensus sequence obtained from alignment of the majority of selected sequences from the mammalian PTB-selected pool was: GCCUG(Y/G)UGCYYYYCYYYG(Y/G)CCC. This shows that we have selected unique pyrimidine-rich sequences that bind PTB[11]. When we performed iterative selection-amplification for the RNA-binding domain of the *Drosophila* PTB, we enriched CU-rich sequences interrupted by guanosines. Among the high affinity sequences that the *Drosophila* PTB selected was an 84% pyrimidine-rich sequence: GCUUUCCUCUGUCGCCCUUCUUCGUCCCCUG. In fact, this sequence is similar to the pyrimidine-rich sequence present in the alpha-tropomyosin intron which binds with high affinity to and is regulated by the mammalian PTB[15]. We have successfully used this approach repeatedly to study RNA-binding properties and functions splicing regulators and a splicing factor[11, 15, 16]. **Table 1** shows successful examples of RNA-binding proteins for which SELEX was used to identify their preferred or consensus binding site(s).

Third, an in vitro splicing assay, which is based on alternative 3' splice site choice, shows functional relevance of distinct but overlapping RNA-binding specificities of polypyrimidine-tract binding proteins. Whereas an upstream 3' splice site is used by default, addition of the recombinant PTB leads to activation of the alternative or downstream 3' splice site (**Figure 3**). In contrast, addition of recombinant hnRNP C[17] leads to repression of both 3' splice sites. Addition of the recombinant general splicing factor U2AF[65] reverses the hnRNP C1-mediated 3' splice site repression (**Figure 3**) as well as the PTB-mediated effect on downstream 3' splice site activation (data not shown). A simple explanation for these effects is a direct competition between the binding of the general splice factor U2AF[65] and PTB (also called hnRNP I), which preferentially binds to and represses certain 3' splices sites, or between U2AF[65] and hnRNP C, which binds to and represses both 3' splice sites.

FIGURE & TABLE LEGENDS:

Figure 1: Summary of key steps in iterative selection-amplification process (SELEX).

Figure 2: Enrichment of PTB-binding RNAs. Increasing concentration (filled triangles) of recombinant PTB was used with either radiolabeled pool 0 RNA or the selected pool obtained following six rounds of selection and amplification. Positions of unbound RNA and the RNA:protein complex are indicated.

Figure 3: Splice site switching assay validates distinct binding specificities of pyrimidine-binding proteins. (**A**-Top) Schematics of the splicing substrate. The splicing substrate contains a 5' splice site and two alternative 3' splice sites flanking the intron. Rectangles (open, with horizontal lines, and solid) are exons and the line is an intron. (**A-**Bottom) hnRNP C1 represses the upstream 3'

splice site (without activation of the downstream 3' splice site), whereas PTB leads to activation of the downstream 3' splice site. The splicing substrate was incubated in a HeLa cell nuclear extract. The splicing products (shown on the sides) were analyzed using a primer extension assay[18] with splice-junction primers (arrows), which recognize splicing of the common 5' splice site to either the upstream or the downstream 3' splice site. (**B**) Recombinant U2AF^{65} (rU2AF^{65}) reverses the repressive effect of hnRNP C1. Addition of the recombinant hnRNP C1, PTB, or rU2AF^{65} proteins to the splicing reaction is indicated by the + symbols.

Table 1: Preferred binding sites for some RNA-binding proteins.

DISCUSSION:

Nucleic acid-binding proteins are important regulators of animal and plant development. A key requirement for the SELEX procedure is the development of an assay that can be used to separate protein-bound and unbound RNA fractions. In principle, this assay can be an in vitro binding assay such as the filter-binding assay, the gel mobility shift assay, or a matrix binding assay[19] for recombinant proteins, purified proteins, or protein complexes. The assay can also be an enzymatic assay where the precursors and products (or intermediates) can be separated based on size or some other means[20].

While mutagenesis has been widely used to characterize binding sites for proteins, it is laborious, and time consuming and longer sequences are not as easily amenable to saturation mutagenesis. The significance of the iterative binding and amplification approach described here is that not only does it overcome some of the above limitations, it can most importantly identify previously unknown binding sites and provide important information about nucleotide requirements at each position at the same time.

An important consideration for the success of iterative selection-amplification is binding affinity and specificity. Typically, 12 to 15 rounds of selection-amplification are employed and a sequence space of 10^{12} to 10^{15} molecules can be routinely sampled. The progress and eventual success of the selection-amplification protocol can be monitored using a binding assay or direct sequencing, which monitors affinity for or enrichment of specific sequences in intermediate pools, respectively. While the binding assay was traditionally used, advent of the next generation sequencing allows analysis of sequence enrichment in ways not possible by manual Sanger sequencing[14].

A critical step in the success of SELEX is fold enrichment of the desired molecules at each step. The number of cycles required for SELEX varies and depends on several factors. For example, if fold-enrichment of desired or specific sequences is higher in each round, fewer rounds will be sufficient. However, if an assay allows a high proportion of undesired sequences in the bound pool, additional rounds will become necessary to enrich desired RNA sequences. A limitation of the technique or an unintended consequence of the need for additional cycles of selection-amplification that must be kept in mind is the possibility that it might introduce artefacts or enrich sequences that have unrelated properties such as their ability to amplify. Finally, while some applications benefit from the highest affinity binders, for other uses, a balance must be

struck during the selection-amplification process between binding affinity and function because tightest binding sequences might not necessarily be the most functional sequences in biological contexts (e.g., if a sequence is recognized multiple times by different proteins during splicing).

Among the modifications and troubleshooting to improve the procedure, negative selection or counter selection can be employed to increase specificity. Similarly, use of different partitioning protocols, such as the filter binding assay followed by the gel mobility shift assay, can eliminate enrichment of unwanted sequences that bind, for example, to the nitrocellulose filter or a column matrix[21]. Given that proteins-nucleic acid interactions have both specific and non-specific components, buffer conditions such as salt and pH have effects on RNA-protein interactions. Moreover, use of appropriate protein concentration can have a direct effect on retention of strong, weak and non-specific binders. Selection pressure can be increased in successive rounds, for example, by including a competitor RNA, reducing protein concentration, or reducing the time of incubation. Thus, careful considerations and optimizing these parameters can impact the outcome of the SELEX protocol.

Recently, many variations or modifications of the original SELEX protocol have been developed which overcome some of the limitations mentioned above. These include high throughput-SELEX (HT-SELEX), which combines SELEX and massively parallel sequencing[6], RNAcompete, which involves incubation with excess non-random RNA, pull-down of the bound RNA, fluorescent labeling of RNA, and analysis on microarrays[5], RNA Bind-n-Seq, which combines RNA affinity analysis in a quantitative and high throughput fashion[22], and RAPID-SELEX, which shortens the process and includes a non-amplification step[23].

Chemically modified bases have been used to expand the repertoire of the RNA molecules for specific applications[24]. Diagnostics, therapeutics, as well as molecules with catalytic activities are among the many applications (including in medicine) of the selected molecules[25]. Aptamers complement antibody-based protocols and provide excellent tools whose potential, for example, in diagnostics, therapeutics, and other applications, remains to be fully exploited[26–28]. In the future, for example, clinical benefits are among the numerous desired applications, beyond what the first FDA-approved aptamer (Pegaptanib sodium) could deliver for the age-related macular degeneration. The scalable proteomic technology for protein measurements offers a step toward understanding health and diseases[24].

ACKNOWLEDGMENTS:
The author thanks the National Institutes of Health for the past funding.

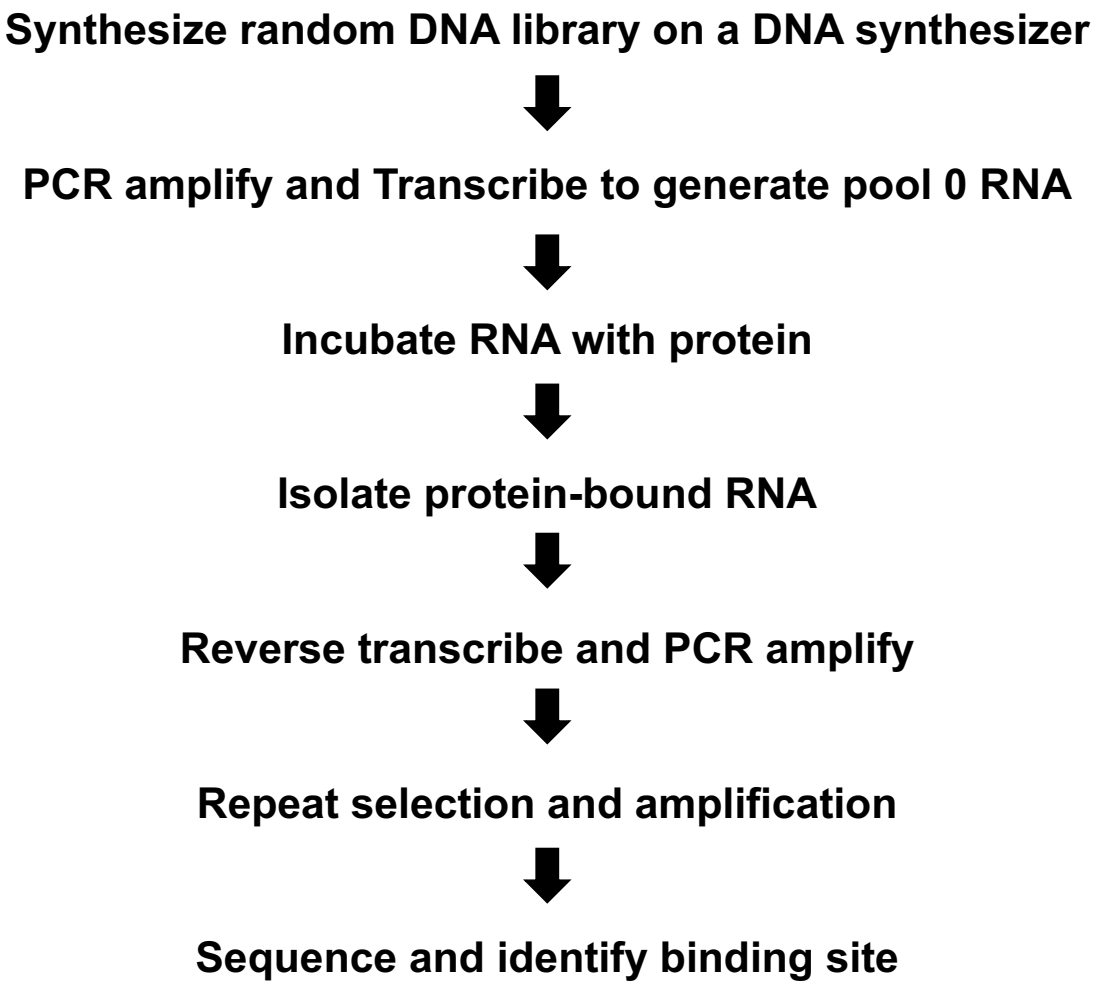
DISCLOSURES:
The author declares that he has no competing financial interests.

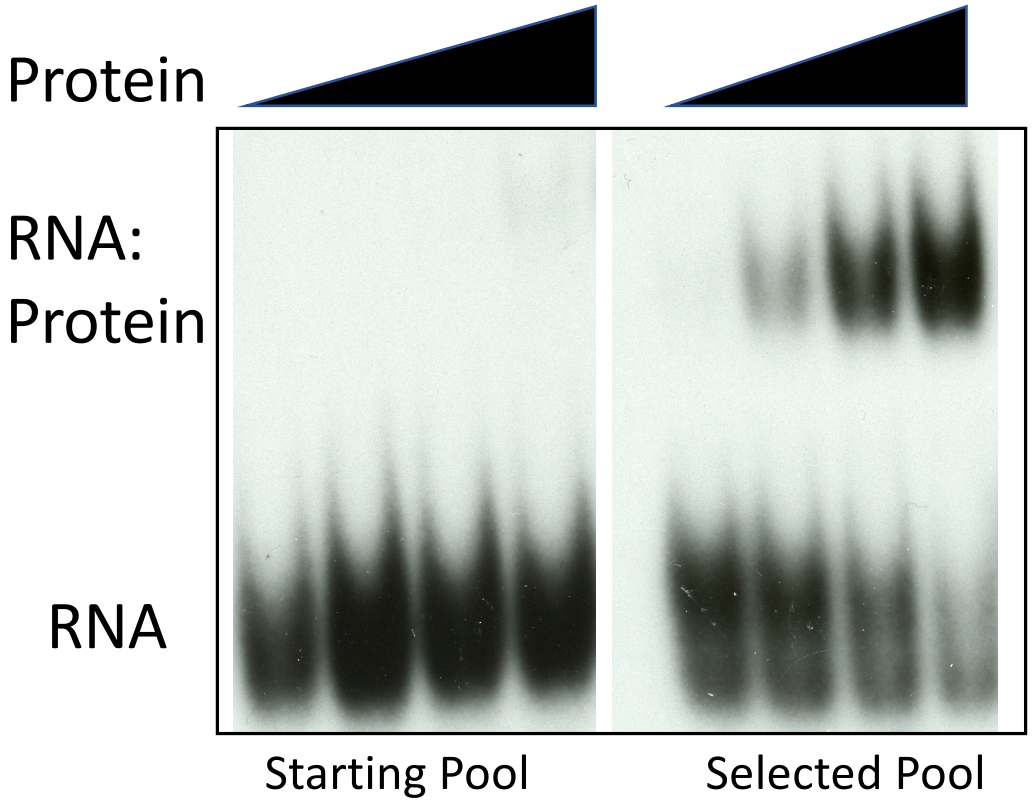
References:

1. Pribnow, D. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences of the United States of America*. **72** (3), 784–8 (1975).

2. Breathnach, R., Chambon, P. Organization and expression of eucaryotic split genes coding for proteins. *Annual Review of Biochemistry*. **50**, 349–83, doi: 10.1146/annurev.bi.50.070181.002025 (1981).
3. Wickens, M., Stephenson, P. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science*. **226** (4678), 1045–51 (1984).
4. Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*. **15** (20), 8125–48 (1987).
5. Ray, D., Ha, K.C.H., Nie, K., Zheng, H., Hughes, T.R., Morris, Q.D. RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. *Methods*. **118–119**, 3–15, doi: 10.1016/j.ymeth.2016.12.003 (2017).
6. Jolma, A., et al. Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. *Genome Research*. **20** (6), 861–73, doi: 10.1101/gr.100552.109 (2010).
7. Tuerk, C., Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. **249** (4968), 505–510 (1990).
8. Ellington, A.D., Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature*. **346** (6287), 818–822 (1990).
9. Cowperthwaite, M.C., Ellington, A.D. Bioinformatic analysis of the contribution of primer sequences to aptamer structures. *Journal of Molecular Evolution*. **67** (1), 95–102, doi: 10.1007/s00239-008-9130-4 (2008).
10. Jenison, R.D., Gill, S.C., Pardi, A., Polisky, B. High-resolution molecular discrimination by RNA. *Science*. **263** (5152), 1425–9 (1994).
11. Singh, R., Valcarcel, J., Green, M.R. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*. **268** (5214), 1173–1176 (1995).
12. Milligan, J.F., Uhlenbeck, O.C. Synthesis of small RNAs using T7 RNA polymerase. *Methods in Enzymology*. **180**, 51–62 (1989).
13. Sambrook, J., Fritsch, E.F., Maniatis, T. *Molecular Cloning*. Cold Spring Harbor Laboratory Press. New York. (1989).
14. Sanger, F., Nicklen, S., Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. **74** (12), 5463–7 (1977).
15. Robida, M., Sridharan, V., Morgan, S., Rao, T., Singh, R. Drosophila polypyrimidine tract-binding protein is necessary for spermatid individualization. *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.1007935107 (2010).
16. Banerjee, H., Rahn, A., Gawande, B., Guth, S., Valcarcel, J., Singh, R. The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF(65)) is essential in vivo but dispensable for activity in vitro. *RNA*. **10** (2), 240–253 (2004).
17. Gorlach, M., Burd, C.G., Dreyfuss, G. The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J Biol Chem*. **269** (37), 23074–8 (1994).
18. Valcarcel, J., Singh, R., Zamore, P.D., Green, M.R. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature*. **362** (6416), 171–175 (1993).
19. Wilson, C., Szostak, J.W. Isolation of a fluorophore-specific DNA aptamer with weak redox activity. *Chemistry & Biology*. **5** (11), 609–17 (1998).

20. Joyce, G.F. Reflections of a Darwinian Engineer. *Journal of Molecular Evolution*. **81** (5–6), 146–9, doi: 10.1007/s00239-015-9724-6 (2015).
21. McKeague, M., Derosa, M.C. Challenges and opportunities for small molecule aptamer development. *Journal of Nucleic Acids*. **2012**, 748913, doi: 10.1155/2012/748913 (2012).
22. Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P.A., Burge, C.B. RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Molecular Cell*. **54** (5), 887–900, doi: 10.1016/j.molcel.2014.04.016 (2014).
23. Szeto, K., et al. RAPID-SELEX for RNA aptamers. *PLoS One*. **8** (12), e82667, doi: 10.1371/journal.pone.0082667 (2013).
24. Rohloff, J.C., et al. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Molecular Therapy - Nucleic Acids*. **3**, e201, doi: 10.1038/mtna.2014.49 (2014).
25. Gold, L., et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. **5** (12), e15004, doi: 10.1371/journal.pone.0015004 (2010).
26. Zhuo, Z., et al. Recent Advances in SELEX Technology and Aptamer Applications in Biomedicine. *International Journal of Molecular Sciences*. **18** (10), doi: 10.3390/ijms18102142 (2017).
27. Blind, M., Blank, M. Aptamer Selection Technology and Recent Advances. *Molecular Therapy - Nucleic Acids*. **4**, e223, doi: 10.1038/mtna.2014.74 (2015).
28. Jijakli, K., et al. The in vitro selection world. *Methods*. **106**, 3–13, doi: 10.1016/j.ymeth.2016.06.003 (2016).





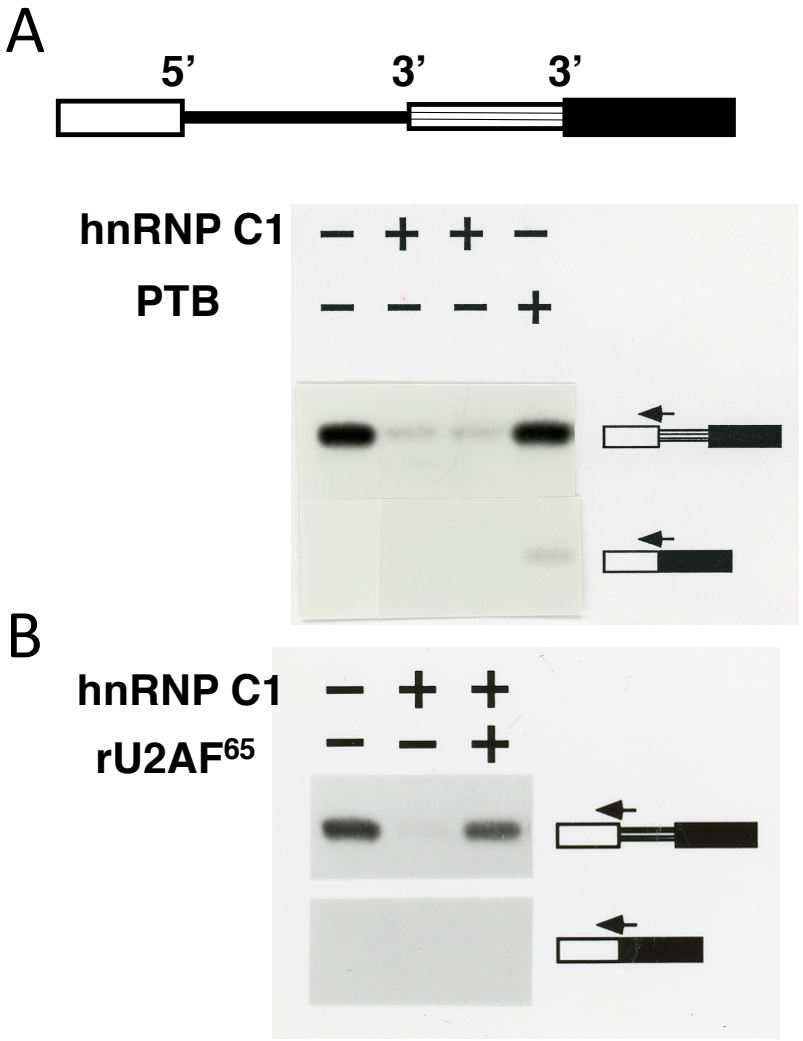


Table 1

Protein	**Preferred sequence(s)**
$U2AF^{65}$	U-rich containing Cs
SXL	U-rich containing 2-4 Gs
PTB	UCUUC-rich with some Gs
hnRNP C1	U-rich (5-6 long)
CstF64	GU-rich
hnRNP E1/E2 and K	C-rich
$U2AF^{65}$/$U2AF^{35}$ heterodimer	UUUYYYYUNUAGGU

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
Gel Electrophoresis equipment	Standard	Standard	
Glass Plates	Standard	Standard	
Nitrocellulose	Millipore	HAWP	
Nitrocellulose	Schleicher & Schuell	PROTRAN	
polyacrylamide gel solutions	Standard	Standard	
Proteinase K	NEB	P8107S	
Recombinant PTB	Laboratory Preparation	Not applicable	
Reverse Transcriptase	NEB	M0277S	
Vacuum manifold	Fisher Scientific	XX1002500	Millipore 25mm Glass Microanalysis Vacuum Filter
Vacuum manifold	Millipore	XX2702552	1225 Sampling Vacuum Manifold
X-ray films	Standard	Standard	

jove 1 Alewife Center #200 Cambridge, MA 02140 tel. 617.945.9051 www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: Exploring Sequence Space to identify binding sites for regulatory RNA-binding proteins

Author(s): RAVINDER SINGH

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☒ Standard Access ☐ Open Access

Item 2: Please select one of the following items:

☒ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	RAVINDER SINGH
Department:	MCD Biology
Institution:	Univ. of Colorado, Boulder, CO 80309, USA
Title:	Assoc. Professor
Signature:	
Date:	Dec. 31, 2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

I sincerely thank the reviewers for their time and valuable input. As a result, the revised manuscripts is significantly improved. I have carefully incorporated reviewers' and editorial suggestions as best as I could. I hope the revised manuscript is suitable for publication.

I have included a line-by-line response below.

Editorial comments:
You will find Editorial comments and Peer-Review comments listed below. Please read this entire email before making edits to your manuscript.
NOTE: Please include a line-by-line response to each of the editorial and reviewer comments in the form of a letter along with the resubmission.

Editorial Comments:

• Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.

• **Textual Overlap:** Significant portions of the manuscript show significant overlap with previously published work. Please re-write the text highlighted in the attached PDF report document (protocol steps 3.8-4.1,4.2-4.3.3,4.4.1-4.4.6) using ALL original text to avoid this overlap.

Re-wrote these sentences in protocol steps 3.8-4.1,4.2-4.3.3,4.4.1-4.4.6.

• **Protocol Detail:** Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. **Please add more specific details (e.g. button clicks for software actions, numerical values for settings, etc) to your protocol steps.** There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol. Some examples:

1) 3.5: Mention centrifuge speed in g, duration and temperature
These details have been added.

2) 8.2: Grow how? Mention incubation conditions and media.
These details have been added.

3) 8.3: Needs a reference
Reference has been added.

4) 9.1: Cite references for offline and online tools.
Reference has been added.

• **Protocol Highlight:** Please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps.
Text has been highlighted

1) The highlighting must include all relevant details that are required to perform the step. For example, if step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be included in the highlighting.
2) The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.
3) Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.
4) Notes cannot be filmed and should be excluded from highlighting.
5) Please bear in mind that software steps without a graphical user interface/calculations/ command line scripting cannot be filmed.

• **Discussion:** JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form (3-6 paragraphs): 1) modifications and troubleshooting, 2) limitations of the technique, 3) significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.
The discussion has been expanded with these details

• Please define all abbreviations at first use.

• Please use standard abbreviations and symbols for SI Units such as µL, mL, L, etc., and abbreviations for non-SI units such as h, min, s for time units. Please use a single space between the numerical value and unit.

• If your figures and tables are original and not published previously or you have already obtained figure permissions, please ignore this comment. If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."

Comments from Peer-Reviewers:

Reviewers' comments:

Reviewer #1:

Manuscript Summary:
RNA-protein binding is central to gene regulation. The author has provided a simple and elegant approach to identify the binding site for known splicing factors.

Major Concerns:
None

Minor Concerns:

The 65 in RNA binding protein U2AF65 should be superscript.
65 is now in superscript throughout the manuscript.

In section 1.1: The author should add that primers as well as random library can be synthesized commercially.
The statement “The primers and the random library can be synthesized commercially” has been added.

In section 3.1: RNA can be transcribed in vitro using commercially available kits with option of T7 or SP6 polymerase.
The statement “RNA can be transcribed *in vitro* using commercially available kits with option of T7 or SP6 polymerase” has been added.

Section 4.3.2: elute RNA from the filter for 2-2 h or overnight with gentle agitation.
This sentence in modified form has been included as part of the re-writing process.

Section 8.1: ligated with appropriately digested pGEM3 or other plasmid carrying restriction sites.
The phrase “or other plasmids carrying restriction sites” has been added.

Section 8.3: Sequencing can be performed in house or commercial source.
The sentence “Sequencing can be performed in house or done commercially” has been added.

Section 9.1: Author may provide link to NCBI or other links for sequence alignment.
A link for multiple sequence alignments options has been added - https://www.ebi.ac.uk/Tools/msa/
.

Figure 3B bottom: The author should clarify the M-tra significance in the contest of this figure.
The label M-tra has been deleted from the figure; M-tra was one of many recombinant intron substrates that we developed. Since it is the only substrate shown it is not necessary to use this name here. Furthermore, description of

this substrate has been provided in the figure legend.

Reviewer #2:

Manuscript Summary:
This is an excellent protocol manuscript describing a powerful selection of RNA sequences that bind to a specific protein. The rationale and experimental design are well described. The protocol is of a broad interest to RNA researchers.

Major Concerns:
How many rounds of selection are often needed? Is there a way to check in the middle rounds that the selection is going well or in the right direction? Is there a simple method to monitor the process (that is, to see if the complexity of the RNA pool is decreasing toward a few "winners")? If yes, perhaps a figure to show the assay results will be good to have.
I have added discussion on the number of rounds and monitoring of the progress of SELEX.

Minor Concerns:
In Fig. 1, (a) PCR amplify, (b) binding sites.
Not sure what to change

In Fig. 3 legend, missing explanation for M-tra.
As noted above, the label M-tra has been deleted from the figure.

Reviewer #3:

Manuscript Summary:
The author provides a brief description of a the widely used SELEX method to uncover RNA binding protein specificity.

Major Concerns:

Overall the introduction and discussion need to be heavily edited. The author should take care to cite previous work on this topic, which has been broadly studied. Key studies using similar approaches such as HT-SELEX, RNAcompete and RNA bind-n-seq were not discussed.
I have provided additional references and expanded both introduction and discussion to include these valuable suggestions.

The author should include a brief description of why certain components are present in buffers (e.g NaCl or DTT). An effort should be made to describe alternative buffers/salts/concentrations and why one would choose to tweak these conditions. While the manuscript is constructed around a single protein, the importance of the technique is that it could be broadly applicable to other factors, thus how one would approach studying other factors should be discussed.
I have included a brief discussion on alternative buffers/salts/concentrations and the need to tweak these conditions.

As this is a methods paper using generic language like "in the freezer" should be avoided. Times, temperatures and volumes should be at the very least suggested.
I have included these details at appropriate places, without excessively repeating details.

Analysis is key to this type of experiment, yet the author spends almost no time describing how one should analyze or interpret the result. For example, how many inserts should be sequenced? Additionally, no primary data of the specific sequences enriched are shown in the example. The author should include examples of these data.
I have included an extra figure with little animation, which specifically includes analysis and interpretation. Furthermore, I have added a Table to expand the examples of successful SELEX and provided consensus sequences for several proteins (rather than specific sequences for just one protein), as examples of these data. This is consistent with this reviewer's sentiment (below) to "avoid simply discussing their previous work" and to reach broader audience for the article.

While I appreciate the discussion regarding the ability of the assay to discern biology, the author should avoid simply discussing their previous work and focus on the caveats, alternative approaches and interpretation of results.

Additional discussion in the revised manuscript should help with some of these concerns. I have deleted or considerably shortened our past work.

Minor Concerns:

Several grammar/typos throughout should be corrected.
I have done both Spell-check and grammar-check.

In the long abstract the following sentence is difficult to understand, please edit accordingly: "These interactions are identified by mutational approaches when the target sequence is somehow localized to a specific gene or region of DNA."
This sentence has been revised.

Please provide citation(s): "Mutational and phylogenetic sequence comparison have been used to identify and subsequently characterize protein-binding sites in nucleic acids. However, numerous orphan proteins are known whose binding site(s) remain to be identified."
Some example references have been added.

The statement needs to be corrected and a short description of available methods needs to be included. A real effort should be made by the author at helping the reader understand current approaches and the pros and cons of each: "However, if the binding site for a protein is unknown, the approach described here is possibly the only strategy to identify it."
Additional discussion (both introduction and discussion) and revision should help with this concern.

Vague statements should be avoided: "Extract with phenol-chloroform, chloroform, and collect aqueous phase. Add Sodium Acetate and ethanol and after incubation in freezer" and "The protein storage buffer provides sufficient glycerol for sample loading into wells." and "Align sequences and obtain a consensus binding site(s) using
Revised text at many places should overcome this concern.

Forward Primer

5’-GTAATACGACTCACTATAGGGTGATCAGATTCTGATCCA-3’

+1

REVERSE Primer

5’-GCGACGGATCCAAGCTTCA-3’

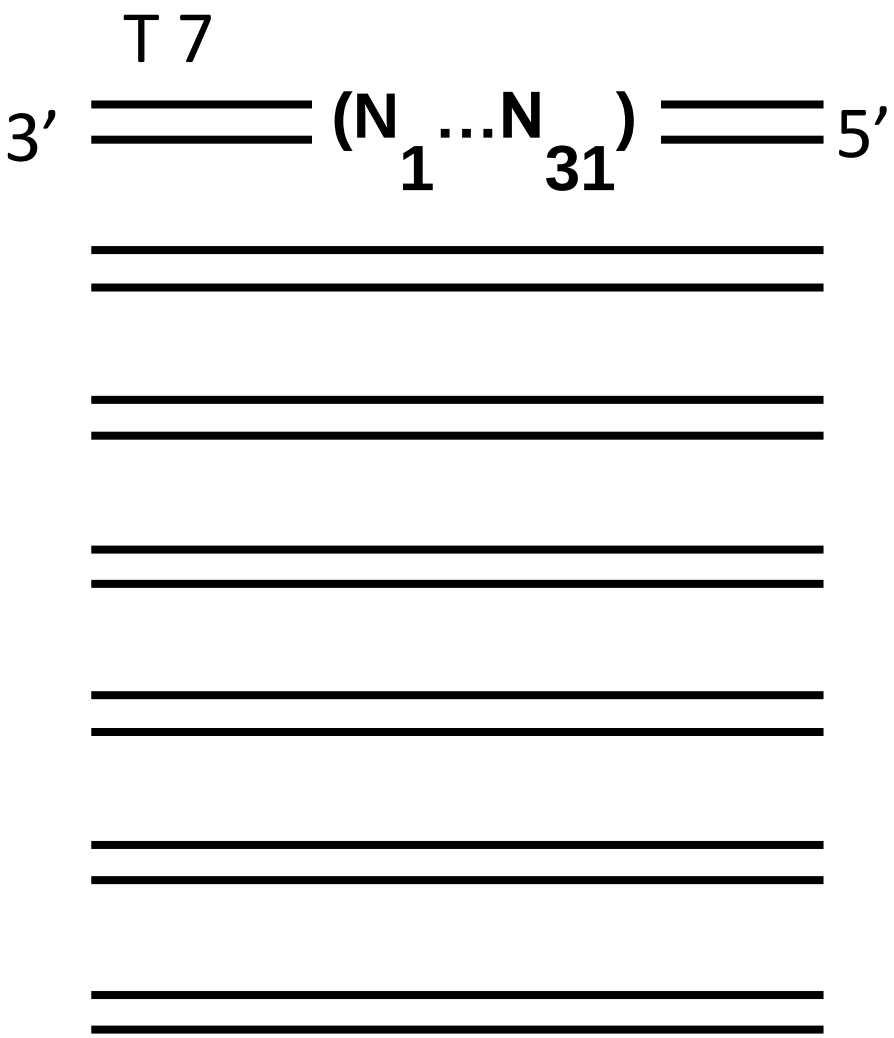
RANDOM LIBRARY TEMPLATE

5’-GGTGATCAGATTCTGATCCA(**N_1…N_{31}**)TGAAGCTTGGATCCGTCGC-3’

Bcl1

HindIII

RANDOM DNA LIBRARY



TRANSCRIPTION

RNA (POOL 0)

5’ *——— (AA ..) ——— 3’

*——— (AG..) ———

*——— (AC..) ———

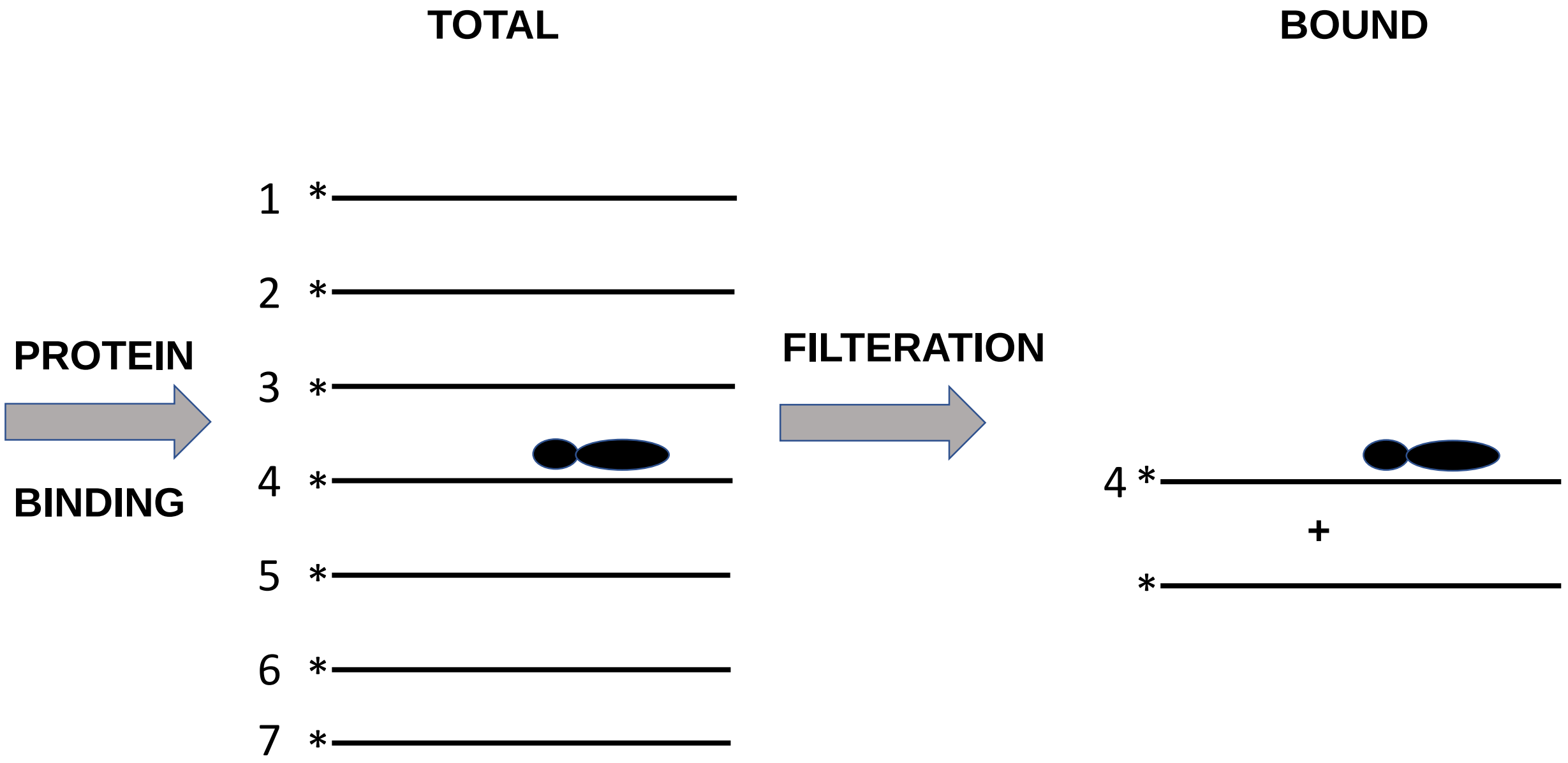
*——— (AU..) ———

*——— (GA..) ———

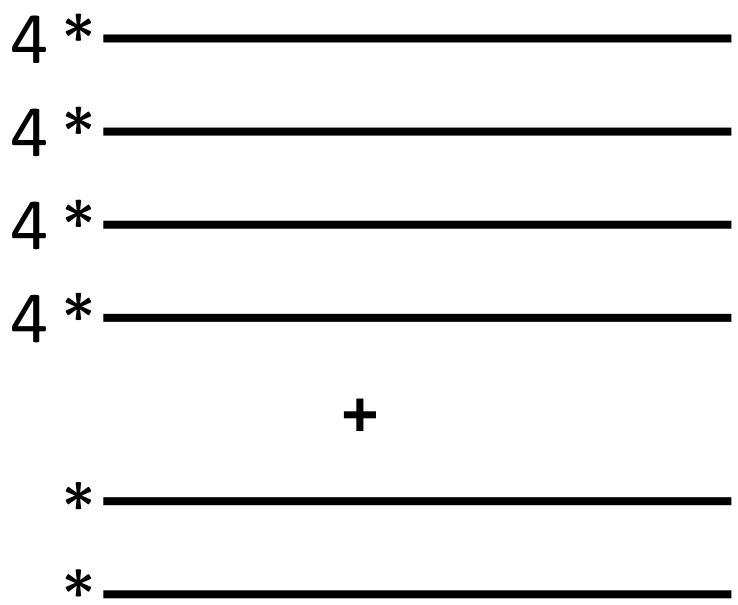
*——— (GG..) ———

*——— (GC..) ———

*——— (GU..) ———


TOTAL
BOUND
PROTEIN
BINDING
FILTERATION
1 *
2 *
3 *
4 *
5 *
6 *
7 *
4 *
+
*

AMPLIFY



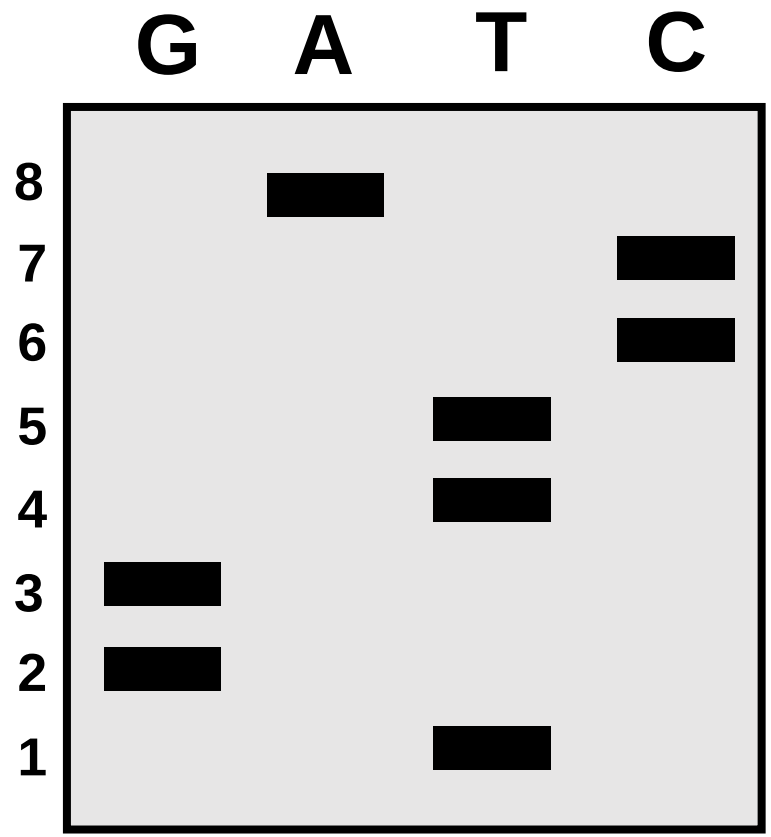
REPEAT

BINDING AMPLIFICATION

CLONE and SEQUENCE ALIGNMENT

CLONE

SEQUENCE



ALIGNMENT

Consensus

NNNYUNNNN

UGGUUCCA
CCGUUCGC
GAGUUUAU
AUGCUAUA
UAGCUGAG
GCGUUGUG
AAGUUAGA
CUGCUUCC