Stochastic Model and Optimization of SELEX

Yue Wang

Department of Computational Medicine
University of California, Los Angeles

Joint work with Bhaven A. Mistry and Tom Chou

May 18, 2022

Outline

- Introduce SELEX: a process to select aptamers.
- Review the traditional deterministic model.
- Build a stochastic model for SELEX and analyze its properties.
- Search for the optimal protocol of SELEX.

Section I: Introduction

- Aptamers are oligonucleotide or peptide molecules that bind to a specific target molecule.
- Some aptamers have high affinity to a pathogen. If we can isolate such aptamers, they can be used as an alternative of antibodies.
- In general, we start with a large library of randomly generated aptamers, and they have different affinities to the target.
- How to select the best aptamers in an easy way?

Section I: Introduction

- Systematic Evolution of Ligands by EXponential enrichment (SELEX): a convenient method to select the best aptamers.
- Aptamers with higher affinities to the targets are more likely to bind to the targets. We can use the targets to pick out such aptamers.
- It is similar to a population evolution process.

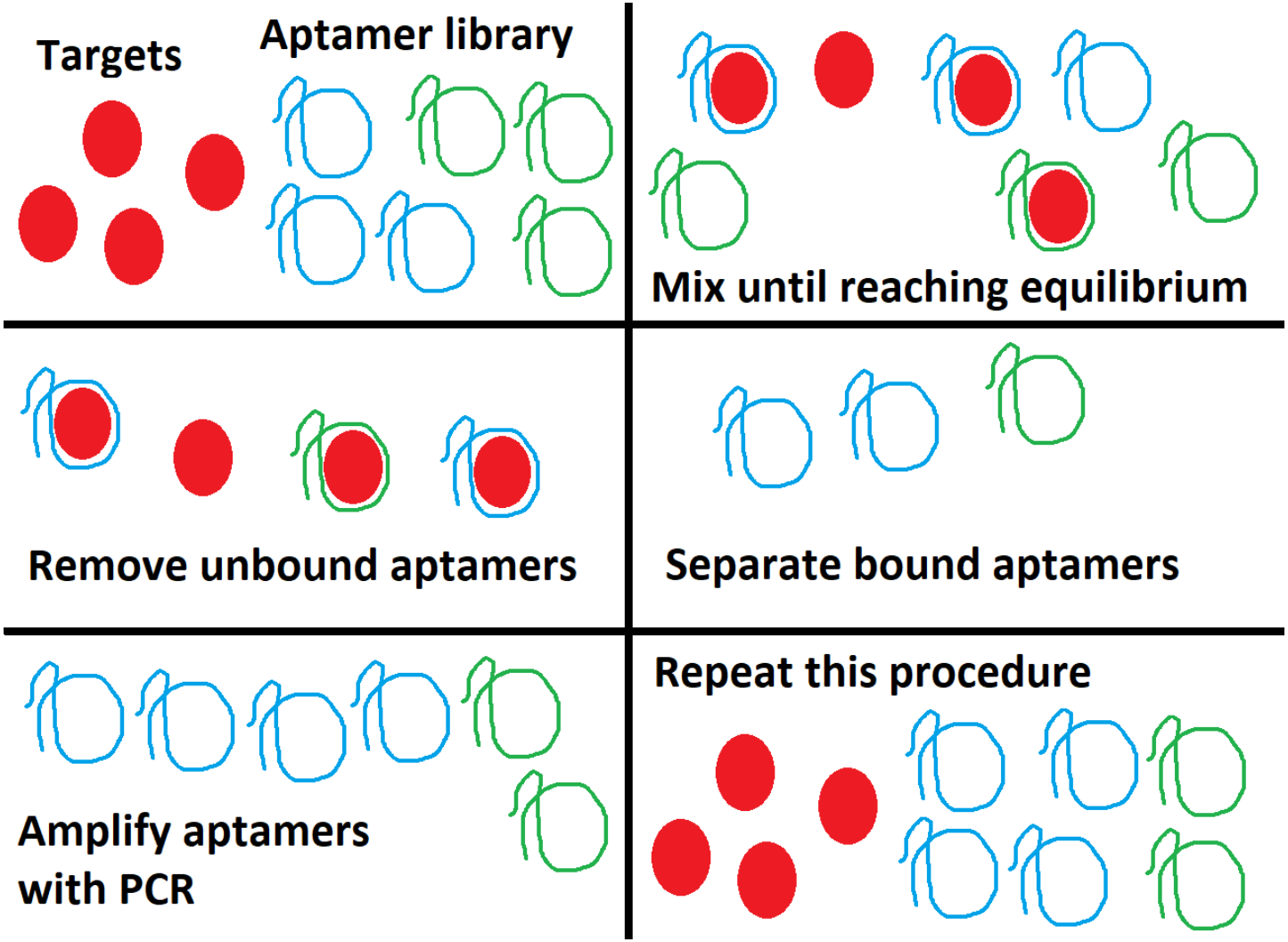


Figure: Protocol of SELEX

Section I: Introduction

- We start with some target molecules and different types of aptamer molecules.
- Mix targets and aptamers until they reach equilibrium. Aptamers with higher affinities are more likely to bind to targets.
- Unbound aptamers are removed. Then we separate bound aptamers.
- Separated aptamers are amplified with PCR.
- One round of SELEX finishes. Aptamers with higher affinities are more concentrated.
- Start more rounds of SELEX to further concentrate aptamers with higher affinities.

Section I: Introduction

- When starting one round of SELEX, we can control the quantity of targets and the quantity of aptamers, but the proportions of different aptamer types cannot be controlled.
- The goal is to maximize the proportion of the best aptamer (with the highest affinity) after this round of SELEX.
- A bad choice: add too many targets, so that almost all aptamers are bound. No selection is made, and the proportion of the best aptamer is invariant.
- We need a mathematical model to study the optimization of SELEX.

Section II: Deterministic Model

- A traditional approach uses the law of mass action, which is valid when the number of molecules is sufficiently large.
- Notations: $[S]$: total concentration of targets; $[A_i]$: total concentration of aptamer type $\mathtt{A}_i$; $[a_i]$: concentration of aptamers $\mathtt{A}_i$ that are bound to targets at equilibrium. K_i^+: reaction constant of $\mathtt{S} + \mathtt{A}_i \rightarrow \mathtt{SA}_i$; K_i^-: reaction constant of $\mathtt{SA}_i \rightarrow \mathtt{S} + \mathtt{A}_i$.
- At stationary, for each reaction $\mathtt{S} + \mathtt{A}_i \rightleftharpoons \mathtt{SA}_i$, we have:

$$\left([S] - \sum_{j=1}^{N} [a_j] \right) ([A_i] - [a_i]) K_i^+ = [a_i] K_i^-.$$

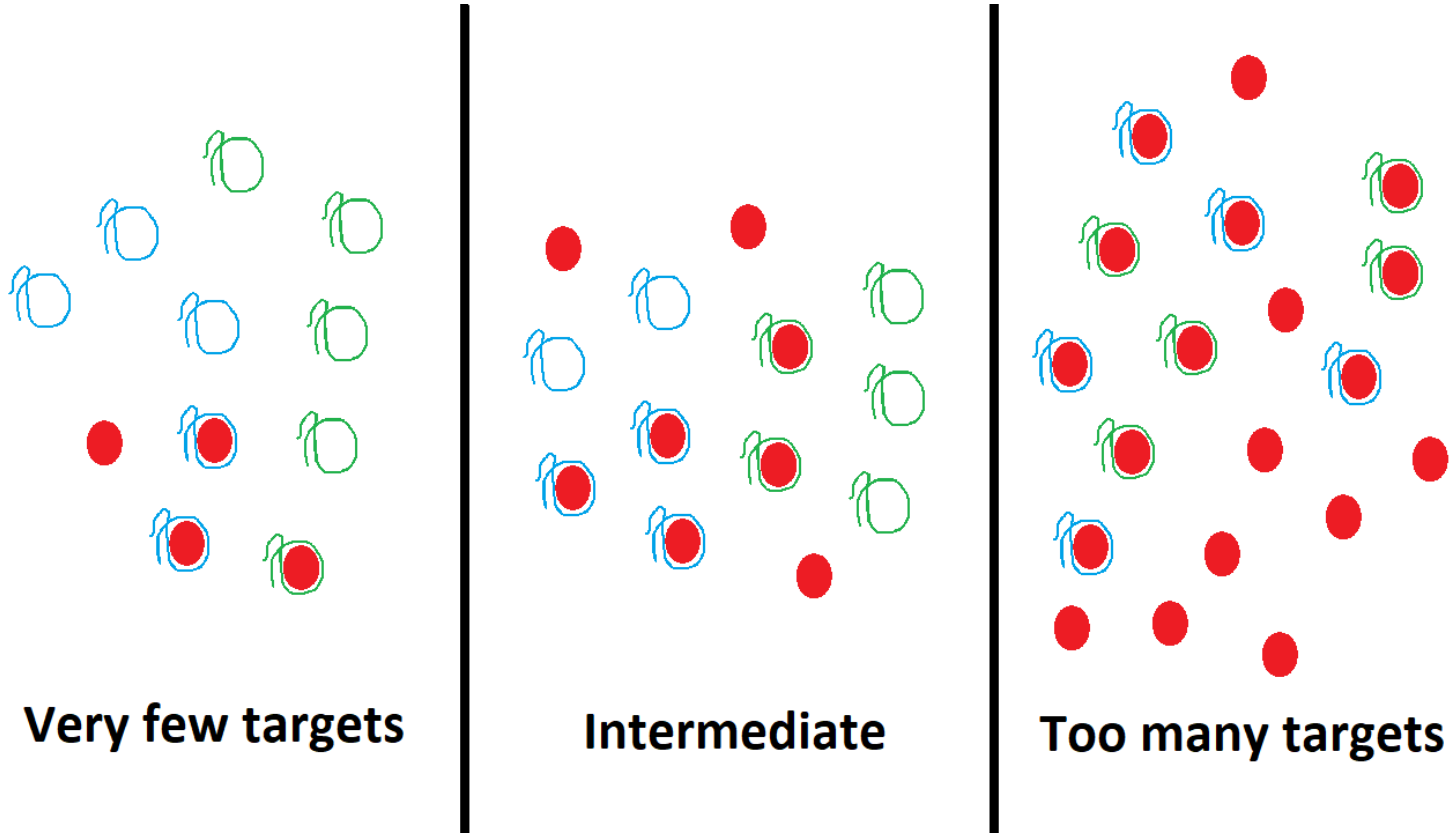
unbound target | unbound aptamer | bound aptamer

- We can use M equations for M aptamer types to solve $[a_i]$.

Section II: Deterministic Model

- Define $K_i = K_i^+ / K_i^-$, and assume $K_1 \geq K_2 \geq \cdots \geq K_M$, so that $\mathtt{A}_1$ is the best aptamer with the highest affinity.
- The goal is to maximize the proportion of bound $\mathtt{A}_1$: $[a_1]/\sum_{i=1}^{M}[a_i]$.
- We can set different values of target concentration $[S]$ and aptamer concentration $[A_i]$, but the ratio $[A_1]/[A_i]$ is fixed.
- $[a_1]/\sum_{i=1}^{M}[a_i]$ increases with $[A_1]$ (and other $[A_i]$), and decreases with $[S]$.
- The optimal policy in the deterministic model: add as many aptamers as possible, and as few targets as possible.

Optimal policy: $[A_i] \gg [S]$.



$$\frac{[a_1]}{[a_1]+[a_2]} = \frac{[A_1]K_1}{[A_1]K_1+[A_2]K_2}$$

upper bound

$$\frac{[a_1]}{[a_1]+[a_2]} = \frac{[A_1]}{[A_1]+[A_2]}$$

lower bound

Section II: Deterministic Model

- The optimal policy in the deterministic model requires very large aptamer concentration $[A_i]$ and very small target concentration $[S]$.
- When $[S]$ is too small, randomness is inevitable, and the law of mass action does not hold.
- We need a stochastic model.
- There is something different in this stochastic model.

Section III: Stochastic Model

- Notations: S: total number of targets; A_i: total number of $\mathbb{A}_i$ type aptamers; a_i: number of $\mathbb{A}_i$ aptamers that are bound to targets. $\bar{K}_i = K_i/V$: reaction coefficient, where V is the system volume.
- Consider a continuous-time Markov process on M-dimensional lattice $\mathbb{Z}^M$, where the states are the bound aptamer counts $(a_1, \ldots, a_i, \ldots, a_M)$.
- The transition rates satisfy

$$\frac{r[(a_1, \ldots, a_i, \ldots, a_M) \to (a_1, \ldots, a_i + 1, \ldots, a_M)]}{r[(a_1, \ldots, a_i + 1, \ldots, a_M) \to (a_1, \ldots, a_i, \ldots, a_M)]}$$

$$= \frac{(S - \sum_{j=1}^{M} a_j)(A_i - a_i)}{a_i + 1} \bar{K}_i.$$

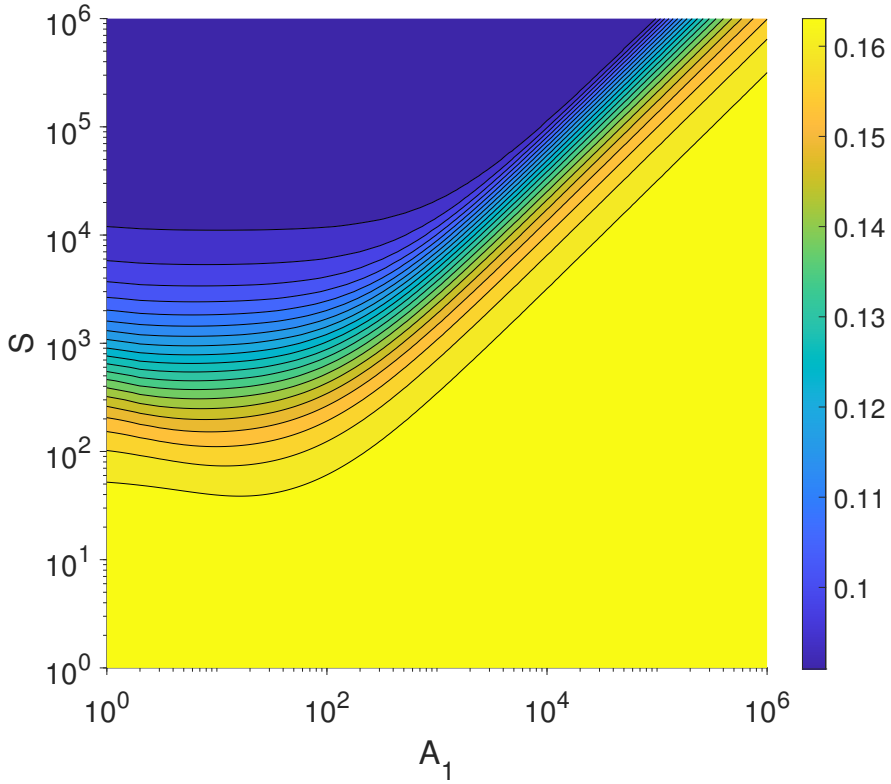
- The stationary probability distribution satisfies

$$\mathbb{P}(a_1,\ldots,a_M) = \mathbb{P}(0,\ldots,0) \times \binom{S}{S-\sum_{j=1}^{M} a_j, a_1, \ldots, a_M} \times \left[\prod_{i=1}^{M} \binom{A_i}{a_i}\right] \times \left[\prod_{i=1}^{M} a_i!\right] \times \left[\prod_{i=1}^{M} \bar{K}_i^{a_i}\right]$$

- For simplicity, we only consider two aptamers $\mathtt{A}_1, \mathtt{A}_2$.
- Due to stochasticity, we need to consider the expected proportion of $\mathtt{A}_1$ and exclude the case that no aptamer is bound: $\mathbb{E}[a_1/(a_1+a_2) \mid a_1 + a_2 > 0]$.

Section III: Stochastic Model

- Contour plot of the $\mathtt{A}_1$ proportion $\mathbb{E}[a_1/(a_1+a_2) \mid a_1+a_2>0]$ as a function of aptamer number A_1 (and A_2) and target number S:



- $\mathtt{A}_1$ proportion decreases with S, but does not always increase with A_1.

In the stochastic model, we still have similar bounds for $\mathtt{A}_1$ proportion.

Theorem

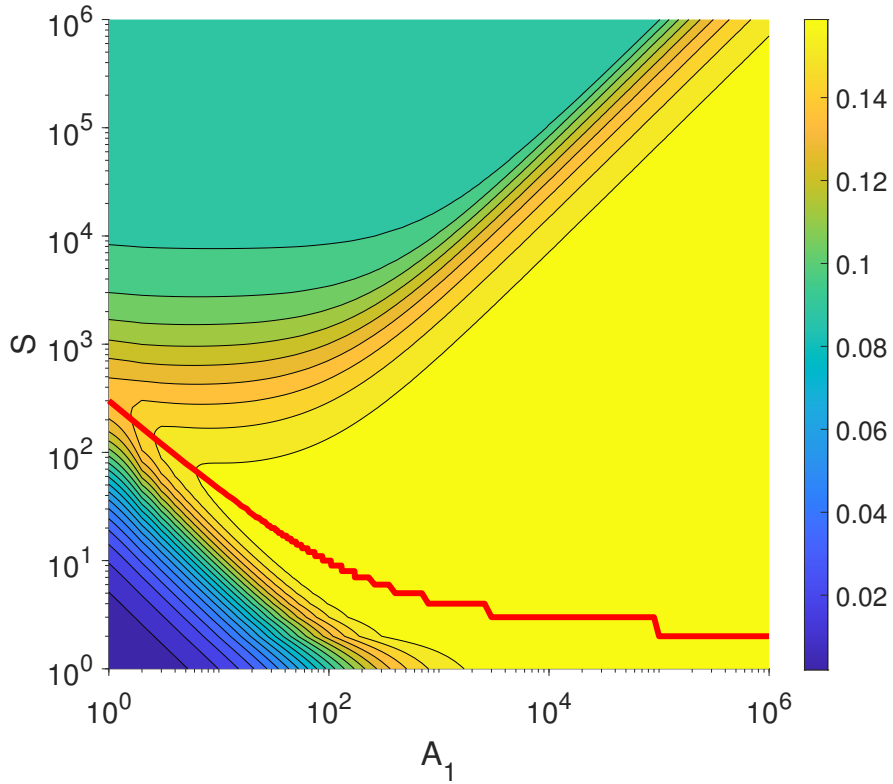
$$\mathbb{E}[a_1/(a_1+a_2) \mid a_1+a_2>0] \leq A_1\bar{K}_1/(A_1\bar{K}_1+A_2\bar{K}_2).$$

Theorem

$$\mathbb{E}[a_1/(a_1+a_2) \mid a_1+a_2>0] \geq A_1/(A_1+A_2).$$

Section III: Stochastic Model

- When $a_1 = a_2 = 0$, stipulate that $a_1/(a_1 + a_2) = 0$.
- $\mathbb{E}[a_1/(a_1+a_2)] = \mathbb{E}[a_1/(a_1+a_2) \mid a_1+a_2 > 0] \times [1-\mathbb{P}(0,0)]$.



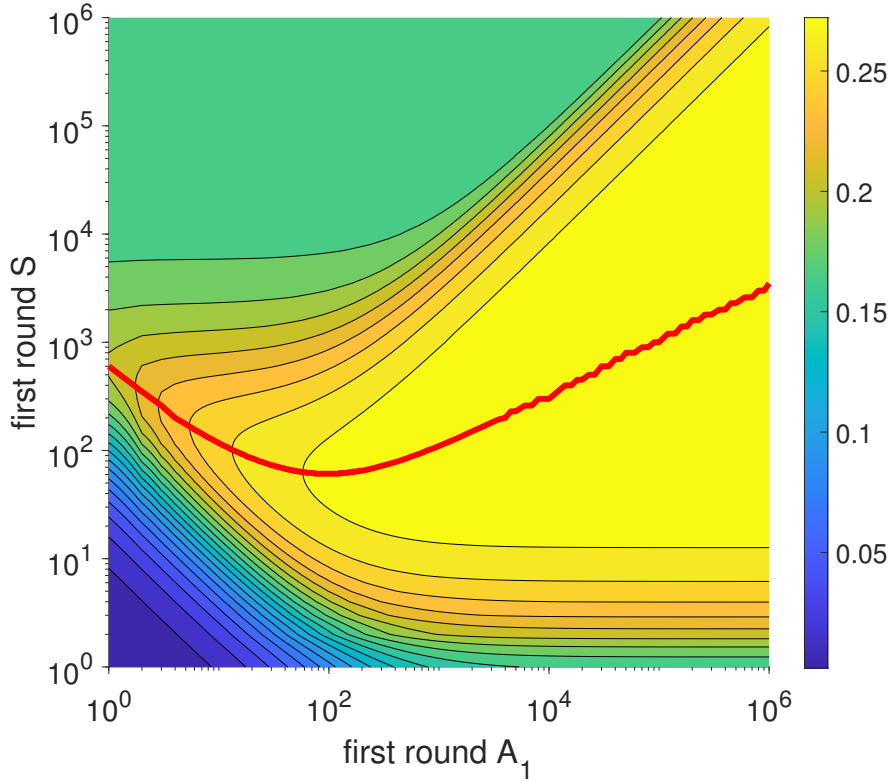
- In the contour plot of $\mathbb{E}[a_1/(a_1 + a_2)]$, when the aptamer number A_1 increases, the optimal target number S (red curve) decreases.

- Optimal policy in the stochastic model:
- When the aptamer number A_1 (and A_2) is not large, the target number S should not too small. Otherwise, $\mathbb{P}(a_1 = 0, a_2 = 0)$ might be large.
- What if we make A_1 (and A_2) sufficiently large, so that $\mathbb{P}(a_1 = 0, a_2 = 0) \approx 0$? Can we set $S = 1$ now?

- For one round of SELEX, $S = 1$ and very large A_1, A_2 can reach the upper bound:
  $\mathbb{E}[a_1/(a_1 + a_2)] \approx A_1\bar{K}_1/(A_1\bar{K}_1 + A_2\bar{K}_2)$.
- However, since there is only one target molecule, after one round of SELEX, only one aptamer type is left.
- After further rounds of SELEX, the expected $\mathbb{A}_1$ proportion does not increase.

Section III: Stochastic Model

- Contour plot of the $\mathtt{A}_1$ proportion $\mathbb{E}[a_1/(a_1+a_2)]$ after two rounds of SELEX:



- For the first round, a policy with large A_1 and very small S does not perform well. The optimal target number S (red curve) first decreases and then increases with A_1.

- Optimal policy for multiple rounds of SELEX in the stochastic model:
- $A_1, A_2 \gg S$ and $S \gg 1$.
- After N rounds of SELEX, this policy has $\mathbb{E}[a_1/(a_1+a_2)] \approx A_1\bar{K}_1^N/(A_1\bar{K}_1^N + A_2\bar{K}_2^N)$.
- Thus $1-\mathbb{E}[a_1/(a_1+a_2)]$ converges to 0 exponentially fast with the rate $\approx \bar{K}_2/\bar{K}_1$. This is the most important factor for the efficiency of multi-round SELEX.

Section IV: Summary

- We discuss SELEX, a process to select the best aptamer for binding a target.
- In the traditional deterministic model, the optimal policy (for any rounds of SELEX) is to have a very large aptamer number and a very small target number.
- We develop a stochastic model, in which the optimal policy for multiple rounds of SELEX is to have a very large aptamer number but a moderate target number.
- This theoretical analysis can be applied to other scenarios, such as selecting drug-resistant cells.

Thank you!