THE APTAMARKER PLATFORM

HOW IT WORKS

There are three innovations behind the success of the Aptamarker platform.

1.) FRELEX selection – The capacity to partition bound from unbound aptamers without immobilizing targets.

2.) Neomer libraries – The invention of reduced random nucleotide libraries with high levels of structural diversity. This enables the application of 100,000 copies of each of 16.8 million aptamers to all samples (closed sequence space).

3.) Selection of structures – The prediction of the structures of all 16.8 million aptamers and selection based on the structural level. This enables statistically robust and sensitive characterization of selection.

The FRELEX Selection

There is a need to partition aptamers that bind to a target in a sample from aptamers that do not bind. We have achieved this with what we call FRELEX selection. We have designed an antisense oligonucleotide that will hybridize to all sequences in an Aptamarker library. The hybridization is designed to occur at three separate points in the library sequence thus ensuring that both hybridization and target binding cannot occur simultaneously.

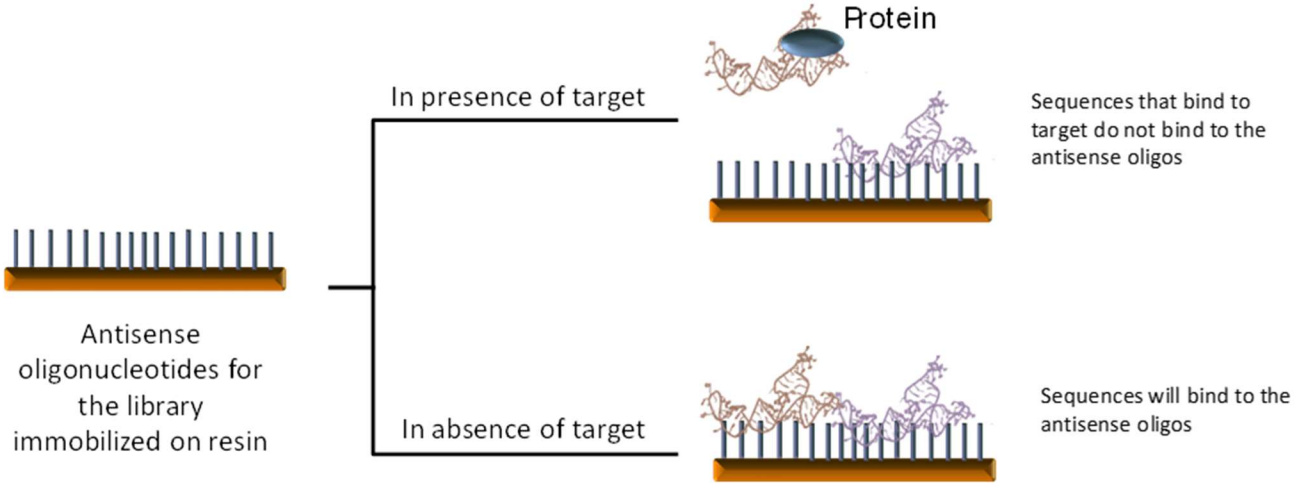

Figure 1: Schematic illustration of FRELEX process

The antisense is immobilized onto the surface of a resin that can be removed from the solution by microcentrifugation. This added to a mixture of the Aptamarker library in a sample. If a given sequence binds to a target molecule then it will not hybridize to the antisense. The mixture is microcentrifuged and the supernatant containing Aptamarkers bound to a target is processed for NGS analysis. There will be some sequences for every Aptamarker that remain in the supernatant without binding to a target, but the abundance of the Aptamarkers in the supernatant will be directly and linearly correlated to the abundance of the target that they bind to.

Closed sequence library

Aptamer development by others including Somalogic is based on the SELEX approach. Key to SELEX is the design of an aptamer library composed of a contiguous random region of 40 nucleotides (nt) flanked by known primer recognition sequences.

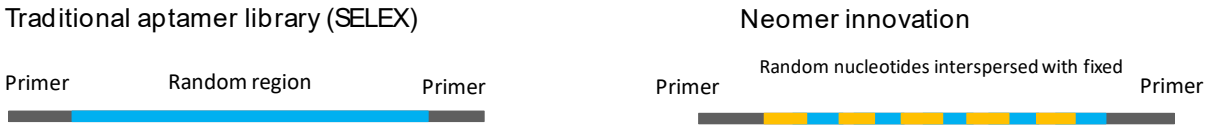

Figure 2: Schematic comparison of SELEX library design with a Neomer library.

The contiguous random region of 40 nucleotides in the SELEX design means that there are 4^{40} = 1.2 x 10^{24} possible sequences. This is far too many sequences to be effectively exposed to a target, thus it is common to use an aliquot of 1 x 10^{15} sequences. There are two problems with this. One is that this is such a small sampling of the possible solution space that the average copy number per sequence will be one. This means that the selection process needs to be reiterated with PCR amplification rounds each time on average at least ten times. This also means that the majority of the sequences will be arbitrarily lost in the first selection round, when each sequence is represented by only one molecule. The other problem is that every aliquot of the library will be composed of different sequences. It is not possible with SELEX to apply the same initial sequences to different samples.

To overcome this problem, we designed the Neomer library containing only 12 random nucleotides interspersed with fixed sequences. The key to this innovation was our use of meta-analysis to characterize thousands of different templates of the fixed sequence with millions of iterations of each template to identify templates that provide both high levels of structural diversity and a high distribution of sequences per structure.

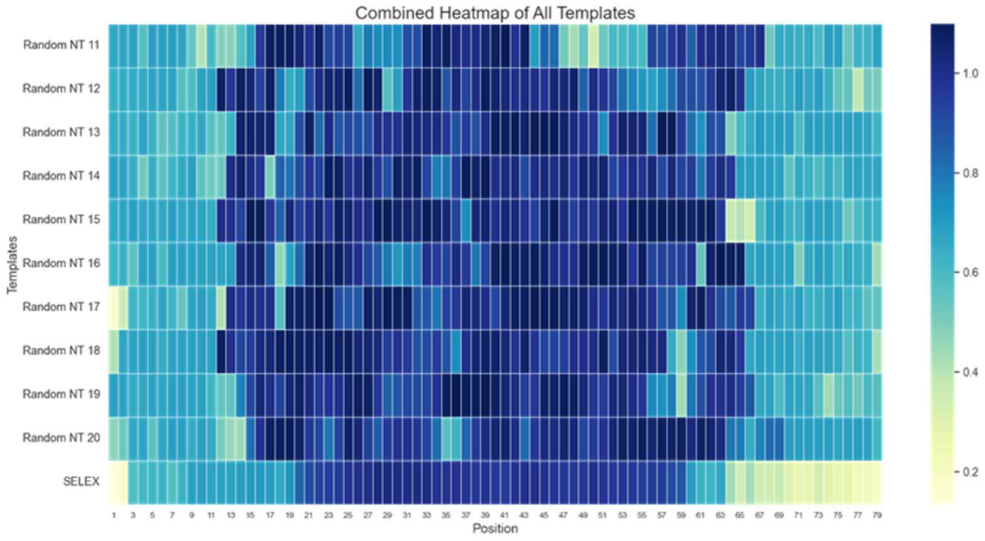

Figure 3: Shannon diversity index characterization of Neomer libraries selected using meta-analysis for structural diversity.

By reducing the number of random nucleotides to 12 we have reduced the number of possible sequences in the library to 16.8M. We are able to apply an average of 100,000 copies of each of the same 16.8 million sequences to different samples. This enables a reduction to one round of selection in order to characterize the effect of each sample on the relative abundance of each of the sequences.

Structure analysis

The reduction of the total number of possible sequences to 16.8M was a tremendous achievement and it enabled reproducible Aptamarker identification across samples. However, 16.8M is still a significantly large number for characterization by NGS analysis. We overcame this constraint by determining the structure of all 16.8M sequences in a library. In the library we are currently using we have identified 130,000 unique structures. This means that with an NGS library read of 10M sequences we have an average of 75 reads per structure. This allows us to detect small changes in protein abundance sensitively and reliably with multiple reads per structure.

The transformation of selection from the sequence level to the structural level allows us to compare selected structures. We are able to identify specific groups of Aptamarkers that are similar in structure that are all binding to the same protein target and groups with a different structure that are also binding to the same protein target. This ability to parse the Aptamarkers that bind to the same target on the basis of their structure enables us to identify different epitopes on proteins and thus discover non-canonical changes such as post-translational modifications, cleavage events, mis-folding, complex formation and isoforms.

The following figure illustrates the binding of this neomer library to Interleukin-6.

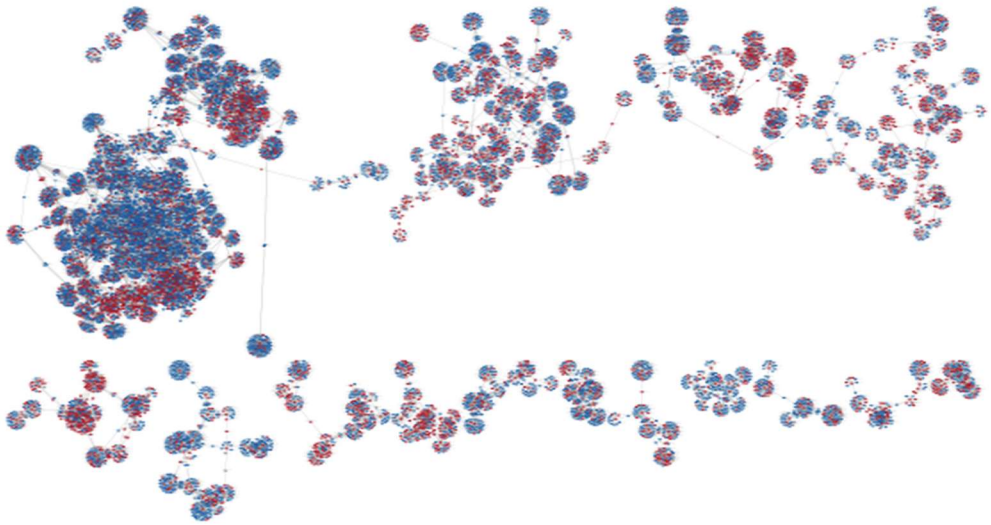

Figure 4: Epitope definition of Aptamarkers for interleukin 6.

Each network of balls represents a potentially different epitope on the target protein. The colouring of the structures from blue to red represents a heat map of their binding strength to this target.

Strategic Application of Aptamarkers:

In existing projects we are proceeding with the Aptamarker library in two simultaneous and complementary approaches.

Agnostic:

We apply the Aptamarker library to samples that correspond to a difference in a disease state. This could be healthy versus affected by a disease, or different stages of the same disease, or disease with and without a specific comorbidity. The nature of the contrast is unlimited. The quality of our results will depend however on the quality of the contrast. We identify which Aptamarkers in our library are binding to biomarkers for the trait in question through next generation sequencing analysis of the selected libraries. The performance of candidate Aptamarkers can be confirmed through qPCR analysis.

We also use machine learning to identify patterns of Aptamarkers that will predict the disease state of interest.

Defined:

We apply the same library to individual candidate biomarkers for the contrast of interest. We use this approach to identify the pair-wise relationship between Aptamarkers and epitopes in your protein. Epitopes differences can be confirmed with wet-lab competition assays on surface plasmon resonance. These defined Aptamarkers are then characterized in the agnostic library selection described above. If there are no post-translational differences in the biomarkers across samples then the defined Aptamarkers should exhibit a high level of covariance amongst themselves. If there are changes to a specific epitope, then the Aptamarker binding to this epitope will lose covariance. We can use this to refine the predictive models built with the agnostic approach above or built previously in your laboratory with Olink or Somascan data on the same samples.

Summary:

The Aptamarker platform is simple to apply and has the capacity to provide results beyond what is possible with other High-Plex proteomics approaches. The beauty and power of this platform is in its simplicity and reproducibility.

To learn more about how you can move your research to the cutting edge of biomarker discovery please contact us at info@neoventures-eu.com for a free consultation.