Chapter 2
Cell-SELEX: Aptamer Selection Against Whole Cells

Dihua Shangguan, Tao Bing and Nan Zhang

Abstract Changes at the molecular level always occur at different stages in diseased cells. The detection of these changes is critical for understanding the molecular mechanisms underlying pathogenesis, as well as accurately diagnosing disease states and monitoring therapeutic modalities. Cell-SELEX is a foundational tool used to select probes able to recognize molecular signatures on the surface of diseased cells. This technology has been increasingly used in biomarker discovery, as well as cancer diagnosis and therapy. In this chapter, the whole cell-SELEX process is described, including aptamer selection, identification, and validation. In addition, we will explore the challenges and prospects for cell-SELEX now and in the coming years. It is anticipated that this chapter will guide readers toward a better understanding of the working principles underlying the cell-SELEX technology and serve as a practical reference for bench scientists engaged in cell and molecular biology.

Keywords Aptamers · Cell-SELEX · Molecular probes · Biomarkers · Molecular recognition

2.1 Introduction

Revealing the molecular mechanisms that trigger changes in biological cells is the concern of scientists across a broad spectrum of disciplines. However, the detection of such changes and, hence, their understanding is primarily thwarted by the lack of tools able to recognize features of cellular architecture at the molecular level.
For cancer diagnosis, pathologists commonly use morphological evidence as a basis for diagnosis, such as anatomical features identified through the use of microscopy, for example, cannot reveal changes, including mass-density fluctuations, at the molecular level, even though such data could be critical to the diagnosis of early stage cancer. Indeed, cancer diagnosis based on data collected through the use of proteomic technologies provide the link between genes, proteins, and disease. Current technologies do not provide for the detection of a cancer cell’s particular molecular signatures, as most methodologies rely on known biomarkers for the development of corresponding probes. However, the number of biomarkers that have thus far been identified and validated is too small to clearly identify even one particular cancer. Besides cancers, the lack of effective probes and biomarkers is also the challenge of the molecular diagnosis of many other diseases. For example, the infectious diseases are extremely widespread all over the world. Because few biomarkers are currently known and available for effective detection of viruses, bacteria, and fungi, as well as infected cells; the diagnosis of infectious agents is cumbersome. Therefore, new technologies that can be applied to the discovery of unknown molecular features of diseased cells and pathogens are in demand.

Aptamers are single-stranded DNA (ss-DNA) or RNA oligonucleotides, which have the ability to bind to other molecules with high affinity and specificity. Compared with antibodies, aptamers have a unique repertoire of merits, including, for example, ease of chemical synthesis, high chemical stability, low molecular weight, lack of immunogenicity, and ease of modification and manipulation [1]. These characteristics make aptamers good candidates as effective probes for the recognition of molecular signatures and as target cell-specific ligands for therapeutic purposes. Aptamers are evolved from a random oligonucleotide library by repetitive binding of the oligonucleotides to target molecules by a process known as Systematic Evolution of Ligands by EXponential enrichment (SELEX) [2, 3]. In the early stage of aptamer development, aptamers were generated against simple targets, such as small molecules and purified proteins [4]. From 2001 to 2003, aptamer selection against complex targets, such as red blood cell membranes [5] and whole cells [6–8], was demonstrated. However, at that time, the application of aptamers toward cancer detection was limited by the absence of aptamers able to bind target cancer cells, or, more specifically, target proteins on the cell surface.

By 2003, only a few cancer biomarkers had been identified, and there were either known biomarkers, were very expensive, or were not commercially available. Therefore, Tan group at University of Florida conceived of the idea of generating specific aptamers by using whole cancer cells as targets. Differences at the molecular level between any two given cell types, such as normal versus cancer cells, or different types of cancer cells, represent the molecular signatures of a specific type of cancer cells. Therefore, the ability to obtain aptamers able to distinguish one type of cancer cell from another by the identification of particular
molecular signatures would constitute aptamers also able to be used as molecular probes for cancer identification. This approach, now known as cell-based SELEX or cell-SELEX, conveniently circumvents the limitations previously noted, and a cell-SELEX protocol was developed using the human acute lymphoblastic leukemia cell line, CCRF-CEM (T cell line), as target cells, as well as human diffuse large cell lymphoma cell line, Ramos (B-cell line), as control cells. This resulted in a panel of aptamers that could specifically bind target cancer cells [9], including one able to bind to a membrane protein on the surface of CCRF-CEM cells, protein tyrosine kinase 7, which has been identified as a biomarker for leukemia [10, 11]. Since then, many aptamers have been generated by cell-SELEX [12–22], their targets are ranged from cancer cells, virus-infected cells to bacteria. These aptamers have shown their utility in cell capture, detection and imaging, even in vivo cancer imaging [1].

Thus, cell-SELEX offers the following advantages:

1. Prior knowledge about the molecular features of target cells is unnecessary since the cell-SELEX specifically generates aptamers that can recognize and differentiate the molecular signatures found on a range of abnormal cell types.
2. Countless molecules, especially proteins, are found on the cell surface, and the molecular differences between two cells typically relate to a series of molecules. In cell-SELEX, each of these molecules is a potential target. A successful selection will therefore generate a panel of aptamers for many different targets, and, as a result, such panel of aptamer probes will provide more data for accurate disease diagnosis and, hence, new opportunities for personalized medicine.
3. Aptamers bind to the native state of target molecules, making it possible for aptamer probes, through ligand binding, to directly recognize their cognate target, creating, in turn, a true molecular profile of diseased cells. In addition, the target molecules are naturally anchored on cell surface, so that the bound aptamers can be easily partitioned from the unbound oligonucleotides by centrifuge or wash (for adherent cells) during the SELEX process; it is unnecessary to purify the target molecules and fix them on solid supports.
4. Providing the opportunity to discover new biomarker. Pathological or physiological changes are complex processes that involve many molecular-level changes on cells. Many of these changes are unknown. Cell-SELEX provides the opportunity to generate aptamers that bind to unknown biomarkers. Then, the obtained aptamers can be used, through affinity separation, to purify and identify their targets. These targets have the potential to be new biomarkers.

Based on these advantages, nowadays cell-SELEX technology has been widely used all over the world and a large number of aptamers specific to a variety of cells have been reported. Nonetheless, the whole cell-SELEX procedure involves
multiple complex steps; and the beginners often fail to obtain desired aptamers. In this chapter, we present an overview for the development of DNA aptamers against different types of mammalian cells (cancer cells as example) using cell-SELEX technology. The discussions mainly focus on the key considerations in each step of cell-SELEX procedure. The challenges and prospects of cell-SELEX are also discussed.

### 2.2 Overview of Cell-SELEX Procedure

The SELEX strategy was described primarily in 1990 by Gold and Szostak [2, 3], and then it has been modified over the years in different ways [23]. Briefly, the general process of SELEX involves the incubation of the target of interest with an oligonucleotide library (DNA or RNA), separation of the oligonucleotides-target complexes from the unbound sequences, and amplification of the bound sequences by PCR or RT-PCR to obtain an enriched pool for next round of selection. This process is repeated until the pool is highly enriched for sequences that specifically recognize the target. The enriched pool is then cloned into bacteria and sequenced to obtain the individual sequences. Representatives of these sequences are chemically synthesized, labeled with reporters and tested against the target to determine potential aptamer candidates. In this whole procedure, the most critical step is the partition of the target-bound sequences from unbound sequences, especially for the SELEX using purified target molecules; thus, many modified SELEX strategies have been proposed to simplify this step or to enhance the efficiency of partition [23]. Compared with the SELEX for purified target molecules, the partition step of cell-SELEX is relatively simple, because the unbound sequences can be easily removed by centrifuge or wash (for adherent cells). But after obtaining aptamers using cell-SELEX, a target identification step is necessary.

The typical cell-SELEX procedure is shown in Fig. 2.1. The starting point of a cell-SELEX process is the preparation of a synthesized random oligonucleotide library and the growth of cells of interest, which is discussed in details in Sect. 2.3.

The iterative cycles of cell-SELEX process includes the following steps: incubation of target cells with randomized DNA library or the enriched DNA pool; collection of cells bound with oligonucleotides; elution of bound oligonucleotides on target cells; amplification of eluted oligonucleotides and preparation of enriched oligonucleotide pool; and counter selection (also named subtractive selection) using control cells. The counter selection step is strongly recommended to reduce the nonspecifically binding oligonucleotides and the oligonucleotides that bind common surface molecules present on both types of cells, which can increase the specificity of enriched pool to the target cells. After each or several rounds of selection, a binding assay of the enriched pool to target cells and control cells is performed to monitor the progress of aptamer enrichment. In this iterative process, each round of selection is not a simple repeat of the previous round. In order to obtain aptamers with high affinity and specificity, the pressure of selection is
progressively increased in the course of a SELEX process by modulating the cell number and the binding/washing conditions in later SELEX rounds. For more details concerning the iterative selection process, see Sect. 2.4.

The number of rounds necessary depends on the progress of enrichment (in general 10–20 rounds). If the binding assay shows that the enriched pool has enough affinity and specificity, the PCR products of the last selected pool are cloned and sequenced to obtain the individual sequences. Representative sequences are chosen, synthesized, and applied to binding assays. Finally, the aptamers with high affinity and specificity to target cells can be further optimized and modified for different applications. The details concerning the aptamer identification are shown in Sect. 2.5.

The efficiency of a cell-SELEX depends on many factors, such as the library, the nature and growth state of cells, the selection conditions, the desired target molecules on cell surface, and operation skills. There is no standardized cell-SELEX protocol for any type of cells and purpose. Therefore in the following sections, we do not describe a detailed cell-SELEX protocol (please see our previous paper for
detailed protocol [24]), but give the general principles of each selection step, including library design, cell preparation, choice of selection conditions, enrichment monitoring, and aptamer identification. The key considerations for each step are discussed.

2.3 Preparation of Oligonucleotide Library and Cells

2.3.1 Oligonucleotide Library and Primers

The design and chemical synthesis of a random oligonucleotide library is the start point of any SELEX. This library consists of huge numbers of different ss-DNA sequences that have a central random region (20–80 nt) flanked on either side by constant sequences for primer binding during PCR amplification. For DNA aptamer selection, this library can be directly used for Cell-SELEX. For RNA aptamer selection, the random DNA library has to be converted into a double-stranded DNA (ds-DNA) library by PCR with a sense primer containing the T7 promoter sequence and an antisense primer, and then transformed into a RNA library by T7 RNA polymerase. Since DNA is more stable than RNA, most of cell-SELEX has been performed with DNA oligonucleotide library. In this chapter, we mainly discuss the DNA aptamer selection by cell-SELEX.

Since PCR amplification is an essential step in each selection cycle, a set of DNA sequences including library and primers needs to be prepared before cell-SELEX. An example of primer-library set for cell-SELEX is shown below [9]:

Sense primer: 5′-(dye)-ATACCAGCTTATTCAATT-3′
Antisense primer: 5′-(biotin)-AGATTGCACTTACTATCT-3′
Library: 5′-ATACCAGCTTATTCAATT-N_{52}-AGATAGTAAGTGCAATCT-3′

As shown in above example, the primer sequences and the length of the randomized region are needed to be designed before the library synthesis. The general rules for conventional primer design also apply to primers for cell-SELEX. A good primer pair should result in high PCR amplification efficiency and low nonspecific amplification. For monitoring the progress of selection, the sense primer is labeled at the 5′-end with a fluorescence dye (such as FAM, TMRA, and Cy5). In order to obtain the ss-DNA pool after PCR amplification, the antisense primer is usually labeled with a biotin at the 5′-end, so that the duplex PCR products can adsorbed on the streptavidin-coated matrix, and then the dye-labeled sense strand (aptamer strand) can be separated from the antisense strand by alkaline denaturation. It is worth to remind that the contamination of primers by trace amount of library would severely disrupt the aptamer selection. Thus, synthesis and purification of primers and library should be performed separately.

The length of the randomized region determines the diversity of the library. For cell-SELEX, the length of randomized region is commonly in the range of
20–80 nt. The short libraries are better manageable, cost-effective in chemical synthesis. However, longer randomized regions give the libraries a greater structural complexity, which is important for cell-SELEX because the molecular targets are numerous and unknown. Therefore, a longer random sequence pool may provide better opportunities for the identification of aptamers [23], but too longer random sequence is not necessary.

The amount of synthesized library used for one cell-SELEX is in the range of 20 pmol–20 nmol, which is equivalent to $10^{13}$–$10^{16}$ random sequences.

### 2.3.2 Choice and Maintenance of Cancer Cells

Aptamers have been reported to be generated against various target cells, such as cells highly expressing specific protein of interest [25], a certain type of cancer cells [9], cancer stem cells [17], adipocyte cells [26] virus-infected cells [19, 21], and bacteria [27, 28]. Most of these aptamers are selected by using live cells, except in one case using fixed cells [8], which may due to that the target molecules on live cells are present in their native state. For cancer cell-SELEX, cultured cancer cell lines are usually used. That is because the periods of cell-SELEX are at least 2–3 months; it is hard to obtain enough live cells with stable performance in this long period by primary cell culture.

Although cell-SELEX without counter selection has been reported able to generate aptamers that broadly recognize common cancer biomarkers [12, 29], more specific aptamers for desired cell lines need to be selected by combining a counter selection step. In order to generate aptamers that only recognize the molecular signatures of the target cells, at least one control cell type is used for counter selection to eliminate the sequences that bind to the common molecules present in both types of cells. The choice of target cells and control cells mainly depends on the purpose of the selection and the future applications of aptamers. In general, two closely related cell types are chosen, such as tumor cells and homologous normal cells, differentiated cells and parental cells, drug resistant cancer cells and drug sensitive cancer cells, virus-infected cells and uninfected parental cells, and antibiotic resistant bacterial-strains and antibiotic sensitive bacterial-strains. In the case of aptamer selection for cancer stem cells, the target and control cells were E-cad$^+$ and E-cad$^-$ DU145 cells, respectively, which were isolated by FACS sorting by gating for the top 10 % of E-cad$^+$ cells and the bottom 10 % E-cad$^-$ cells [17].

Cell culture maintenance is very important in cell-SELEX, because improper cell culturing maintenance may influence the aptamer enrichment, and even lead to failure of aptamer selection. For example, overgrowth of cell culture results in higher rates of cell death, which potentially lead to alteration in cell morphology and protein expression, as well as cause collection of a lot of nonspecific oligonucleotide sequences because of the increased membrane permeability of dead cells. The reduction or elimination of dead cells can significantly enhance the
enrichment of selected pool. Also, the change of cell culture condition and cell growth status may result in the changes in expression levels of some target molecules. The viability of adherent monolayer cells may be not a big problem in selection because most dead cells normally float in the medium, and once the medium is removed, relatively healthy cells are obtained. That notwithstanding, cells must not be allowed to overgrow. But for adherent monolayer cells, the density of cells used for each round of selection should keep consistent because the cell-to-cell connections and expression levels of target molecules would be quite different at different cell density.

There are two ways by which adherent monolayer cells can be used for selection: either directly in cultured dish/flask or as dissociated cells. The direct option may offer a better representation of the cells’ natural environment. Dissociated cells either by treatment with short-term trypsin treatment or non-enzymatic dissociation buffer or by scraping may result in the change of the surface expression, cell death, and cell rupture.

2.4 Aptamer Selection Against Live Cells

The key principle for a successful SELEX is to keep the specific binding sequences and reduces the nonspecific binding sequences as much as possible during the whole SELEX process. The appropriate selection conditions and selection process are essential to achieve this principle.

2.4.1 Binding Conditions

The selection conditions play an important role in the success of cell-SELEX. These conditions include binding buffer, washing buffer, elution buffer, and binding temperature.

For cell-SELEX, the most essential requirement for buffer condition is to keep the cells intact during the binding and washing steps. Thus, the osmolarity, ion concentrations, and pH value of binding buffer and washing buffer must match those of cells. The commonly used buffers are cell culture media, phosphate buffered saline (PBS), and other buffered salt solution (such as Hank’s buffer and Tris-HCl buffer), pH 7.4. In our cases, the washing buffer is PBS plus 4.5 g/L glucose and 5 mM MgCl2. As sequences in initial DNA library or enriched pools may nonspecifically bind to some molecules on cell surface (e.g., electrostatic adsorption), excess other nucleic acid sequences, such as t-RNA, salmon/herring sperm DNA or synthesized oligonucleotides are added to the binding buffer to compete the nonspecifically binding of sequences in library or selected pools. Since the added sequences do not contain the primer binding sites and cannot be amplified by PCR, there is no need to worry about the interference on the selection by these
sequences. In order to reduce the nonspecific sequences that binding to proteins on cell surface and to enhance the utility of aptamers in biological samples or in vivo, albumin (e.g., BSA) or serum is usually added to the binding buffer. Therefore, in our cases, the binding buffer is prepared by adding other nucleic acid sequences (0.1 mg/mL) and bovine serum albumin (BSA, 1 mg/mL) or fetal bovine serum (FBS, 1–10 %) into washing buffer.

The function of elution buffer is to dissociate the potential aptamer sequences from cells. For cell-SELEX, the most commonly used method to recover the sequences binding to target cells is by heating the cell-DNA complex at 95 °C in water or buffer. Compared with other used methods, such as phenol/chloroform extraction, TRIzol extraction and denaturation by 7 M urea, the heating method is simple and efficient. That is because: (i) Elevated temperature will cause denaturing of the cell surface proteins and the folded structure of the DNA, and this will lead to the disruption of the interaction between DNA and protein and the release of DNA from the target protein. (ii) At 95 °C, any DNase that is released after cell disruption at the elevated temperature is inactivated and therefore cannot cause DNA digestion. (iii) After heating, the supernatant containing potential aptamers can be directly used as template for PCR amplification.

The binding temperature depends on the purpose of selection and application of aptamers. In general, 4 °C, room temperature and 37 °C are used in cell-SELEX. However, the higher temperatures such as 37 °C can cause internalization of oligonucleotides into live cells, which may result in the collection of nonspecific sequences because that not all internalizations are caused by the specific binding, such as pinocytosis. In addition, incubation with live cells at 37 °C may increase the probability of DNA digestion by nuclease. Most of our cell-SELEX cases have been performed at 4 °C or on ice. Although the binding affinity of some aptamers selected at 4 °C may decrease at 37 °C, most of the aptamers bind very well at 37 °C [15, 16], especially those with very high affinity. Some of the aptamers generated at 4 °C have been used in various applications at 37 °C [9, 18].

2.4.2 Selection and Counter Selection

As shown in Fig. 2.1, the main body of the cell-SELEX process is the iterative cycles that take most of the time for whole SELEX. Each cycle generally includes steps of target-cell binding, washing, elution, PCR amplification, and enriched pool preparation, as well as a counter selection.

Theoretically, there is only one copy of each sequence in the initial library. Therefore, it is highly possible to lose some of the specific sequences in the first round of selection. When any sequence is lost, it can never be recovered. In order to collect as many specific sequences as possible in the first round of selection, the amount of used target cells should be higher than the subsequent rounds; the incubation time of target cells and initial library should be long enough to let specific sequences have more chance to bind to the target molecules on cells.
In order to avoid loss of specific sequences, the washing strength should be moderate, and the counter selection usually is not performed in the first round. The typical protocol for the first round of selection is as follows: incubate target cells (1–20 × 10^6) with synthesized ss-DNA library (20 pmol–20 nmol, 10^{13}–10^{16} random sequences) in 1–5 mL of binding buffer on ice for 0.5–1 h. After incubation, cells are washed with 0.5–1 mL of washing buffer for 1–3 times. Then, the bound sequences are eluted by heating at 95 °C for 5 min in 300 μL of DNase-free water, and all the eluted sequences are directly applied for PCR amplification. Finally, the PCR products are converted to enriched pool for the second round of selection.

From the second round of selection, a counter selection step is usually added into the selection cycle. The counter selection can be carried out before the target-cell binding step [16] or after elution step [9]. In the former case, the selection cycles start with incubation of control cells (10 million or monolayer in a 60 cm^2 dish/T75 flask) with the DNA pool (200 pmol) in binding buffer for 30 min; the supernatant containing unbound sequences is then incubated with target cells, and the other steps from washing to preparation of enriched pool are similar with those described in the first round. In the latter case, the steps before elution are same with those described in the first round. After that, the bound sequences on target cells need to be eluted by heating at 95 °C for 5 min in 300–500 μL of binding buffer, and then incubated with control cells on ice for 1 h. After centrifuge, the supernatant need to be desalted with NAP 5 column (GE Healthcare) and then applied for PCR amplification; and finally, the PCR products are converted to enriched pool for the second round of selection. In order to effectively eliminate the nonspecific sequences, the control cells used for counter selection should be much more than the target cells.

Because after PCR amplification in the first round of selection, the specific sequences in the enriched pool have many copies, the key consideration from the second round is to effectively eliminate the nonspecific binding sequences and weak binding sequences. In order to obtain aptamers with high affinity and specificity, the selection pressure from second round is gradually enhanced by decreasing the amount of the ss-DNA pool (from 200 to 30 pmol), the incubation time for the target-cell binding (from 60 to 10 min), and the target-cell number (from 10 to 0.5 million); and by increasing the number of washes (from 2 to 5), the volume of wash buffer (from 0.5 to 5 mL).

Since cells are grown in suspension or adhesion on a surface, there are some differences in handling the suspension cells and adherent cells during the SELEX process. For suspension cells, the cells bound with DNA sequences are collected by centrifuge in the binding and washing steps. For adherent cells, cells bound with DNA sequences are collected by removing the supernatant with pipette. But after the washing step, the adherent cells need to be scraped off from the bottom of flask or dish, and then cells are collected by centrifuge and applied for elution. Scraping is used to detach cells because it would not affect the binding of aptamers on cell membrane surfaces. The conventional detaching method by trypsin could not be
used because it was demonstrated that trypsin could cleave the target proteins on the cell surface [9, 16].

During the binding and washing steps, strong vortexing/shaking and high-speed centrifuge should be avoided since that can cause cell breakage and may eventually affect selection. When collecting cells bound with DNA sequences in the binding and washing steps, the residue liquid should be removed as much as possible because a small amount of liquid contains a large amount of unbound sequences. Before each round of election, the DNA library or pool needs to be denatured by heating at 95 °C and cooled on ice to obtain folded ss-DNA.

### 2.4.3 Preparation of Enriched Pool

Compared with other combinatorial chemistry strategy, the most attractive advantage of aptamer selection from nucleic acid library is that nucleic acids can be easily amplified by PCR or in vitro transcription. Unlike the common PCR amplification for a specific template sequence, the DNA library and enriched pool are highly complex. In order to obtain highly efficiency of PCR amplification, an optimization of PCR conditions for library is necessary before SELEX. After PCR amplification of each SELEX cycle, a gel electrophoresis assay is needed to assess the PCR. The PCR contamination should be avoided in the whole SELEX process. The detailed protocol for PCR can be obtained from our previous publication [24].

As the PCR products are ds-DNA, they have to be transformed into a new ss-DNA pool by separating the sense strands from the antisense strands. Several methods have been described in literature for this purpose. Based on our experience, the streptavidin/biotin approach is a very effective way to obtain the sense strand from PCR mixture. The general procedure of streptavidin/biotin approach is as follows: The bound sequences on target cells are amplified by PCR with dye-labeled sense primer and biotin-labeled antisense primers; the PCR products are captured by streptavidin-coated support (such as Sepharose beads and magnetic beads), and then the sense ss-DNA strands are eluted off the beads with 0.1–0.2 M NaOH solution (0.5–1 mL). The eluate is desalted with NAP 5 column (GE Healthcare) and quantified the amount of DNA sequences by absorption at 260 nm. Finally, the desalted solution is dried to obtain the enriched pool for the second round of selection.

### 2.4.4 Enrichment Monitoring and Selection Procedure Adjustment

The selection cycle is a complex process. It is hard to ensure that the aptamer sequences are enriched in every selection cycle. Therefore, the progress of the
selection is needed to be monitored by binding assay at the end of every cycle or every several cycles. Based on the results of binding assay, the selection procedure can be adjusted. The binding assay of the enriched pool in cell-SELEX usually performed using confocal imaging and flow cytometry.

Flow cytometry assay is the best way to monitor the enrichment progress of cell-SELEX, because this method has high sensitivity, good reproducibility, high degree of statistical precision due to the large number of cells measured, quantitative nature of the analysis, and high speed. For the suspension cells, the binding assay can be carried out by incubation of cells with dye-labeled DNA pool in binding buffer. After washing, cells are applied to flow cytometry assay. For adherent cells, before incubation with selected pool, cells need to be detached and suspended in binding buffer. The cell detachment is performed with non-enzyme cell dissociation solution (including EDTA). Then, the detached cells are gently dispersed with pipette, resuspended in binding buffer and passed through a 40 μm strainer to remove the cell clusters which would block the tube of flow cytometer. Finally, the cells are incubated with selected pool, washed, and applied to flow cytometry assay. The concentration of selected pool for cell binding is set in the range of 0.1–1 μM.

Confocal imaging (or fluorescence microscopy imaging) can provide intuitive information of the DNA binding on cells. After incubation with selected pool and washing, adherent cells can be directly observed when they are attaching on the bottom of the dishes; suspension cells can be detected after dropped on a thin glass slide and covered with a coverslip. Although confocal imaging can show the binding regions of DNA sequences on cells, it is difficult to provide the quantitative information with statistical significance for comparing the binding ability of selected pool from different rounds, thus the results of confocal imaging can be used as a supplement to confirm the binding of selected pool on cells. Usually, the selected pools of the first few rounds have low binding affinity, so that the fluorescence intensity on cells is weak, thus the enrichment of the first few round is hard to be judged by confocal imaging.

In order to estimate the enrichment and specificity of the selected pool, controlled experiments with control sequence (unselected library or non-binding sequence) labeled with the same dye and with control cells must be done. The typical progress of a successful cell-SELEX is shown in Fig. 2.2 [15]. The fluorescence intensity of target cells stained by the selected pools gradually enhances with the increase of selection rounds; and the fluorescence intensity of the control cells does not significantly change with the selection going on. These results indicate the steady progress of the selection and the good specificity of the selected pools to target cells. But in practice, the steady progress and good specificity are not often obtained; if so, the SELEX procedure needs to be adjusted.

Cell-SELEX is a highly complex process; many factors can affect the enrichment of aptamer sequences, such as low sequence diversity of DNA library, contamination of PCR reagent, contamination of eluted sequence from target cells (PCR template) by DNA pool or PCR products from previous round of selection, sequence discrimination in PCR, strong nonspecific binding, low abundance of target molecules on cells, and change of cell states.
It often happens that no notable progress is observed after many rounds (5–10) of selection. If so, it is necessary to check for any inappropriate operation during the SELEX. If all the selection steps are run correctly, it may be due to too many nonspecific sequences or too few specific sequences in the selected pool. If that is the case, what we can do is to further enhance the selection pressure as describe in the previous section. In one of our previous case [9], we extra added 20% FBS and 50- to 300-fold molar excess other synthesized DNA in the incubation solution to suppress the nonspecific binding sequences. If all these efforts do not help much, it may be due to the loss of specific sequences in the first round. If so, the SELEX has to be restarted from the first round or even from redesign of the DNA library.

It also often happens that the selected pools bind to both target and control cell lines. It is due to that many of the sequences bound to the common molecules on both cell lines are enriched, and the counter selection is not enough to eliminate these sequences. If that is the case, the counter selection should be strengthened by increasing the number of control cells, reducing the amount of DNA pool, and decreasing the number of target cells. If the enriched sequences bind to highly abundant targets molecules that are present on cell surface of both cell lines, they could not be eliminated only by strengthening the counter selection. In many cases, the counter selection could not completely eliminate the enriched sequences that bind to both cell types. But, if the selected pool bound to target cells stronger than to control cells (Fig. 2.3), aptamers that only bind to target cells may be found in this enriched pool [16].

Fig. 2.2 Monitoring of the cell-SELEX progress. Flow cytometry binding assays of the selected pools from the 4th, 5th, 15th, 18th, and 22nd rounds. Left LoVo cells (target); right HCT-8 cells (control); FITC-labeled ssDNA library as control DNA. Reprinted from Ref. [15], Copyright 2014, with permission from Elsevier.
2.5 Aptamer Identification

2.5.1 Cloning and Sequencing

By iterative cycles of selection and evolution, the complexity of the initial random DNA library is reduced, and aptamer candidates with high affinity and specificity are enriched. The completion of selection is judged by the results of binding assay. The selection circle is finished when the binding assay shows that the selected pool exhibits significant binding ability to target cells, the affinity of the pool cannot be further increased in two or three successive rounds of selection, and the binding of the pool to target cells is higher than that to control cells. Then the final selected pool is PCR-amplified using unmodified primers and cloned into Escherichia coli using the TA cloning kit, and the positive clones (50–100) are sequenced to identify individual aptamer candidates. Since many commercial companies provide cloning and sequencing service and we only need to provide PCR products of the selected pool, the detailed protocol of cloning and sequencing is not described here.

The sequencing result of each clone contains a sense strand or an antisense strand of a potential aptamer. By searching with both primers, the sense and antisense strands of all clones are sought out, and the antisense strands are converted to their complementary strands (sense strands). After removing the primer sequences at both ends, the core sequences (random sequence region) of the potential aptamers are analyzed by sequence alignment using programs, such as Clustal X, clustal omega, and CLUSTAL W. Usually, the sequence alignment groups the potential aptamer sequences in several families (Fig. 2.4). In each family, the sequences are identical or different in only few nucleotide positions. In some cases, special sequence patterns or highly conserved regions are identified among different aptamer groups.
New sequencing technology (454 pyrosequencing) together with bioinformatics (MAFFT 6.0, Perl, Jalview 2.4 and ATV 4.0.5 programs) have also been used to analyze the sequences in enriched pools of different rounds of selection in a large scale [29, 30]. The usage of new sequencing technologies with the help of bioinformatics can provide early and detailed information for progress of SELEX by displaying the intensity of selected sequences in the earlier rounds, as well as provide more information for the choice of candidate aptamers.

2.5.2 Aptamer Candidate Screening and Aptamer Validation

The final aptamer pool may be more complex than we thought, not all the sequences are aptamers. We even have met the situation that most of the obtained sequences did not bind to target cells. After known the sequences of the selected pool, it is necessary to seek out the real aptamers in these sequences. There is no need to synthesize all the obtained sequences because the sequences in a family are
almost identical aptamer. Usually, a representative sequence from each family is chosen for validation. Before chemical synthesis, the flanked primer sequences should be added back to the core region, because the primer sequences may directly participate in the target binding or take part in the formation or maintenance of the binding structure. The full-length sequences are synthesized and labeled with fluorescence dye or other reporters. The binding assay of the synthesized sequences is performed by flow cytometry or confocal imaging. Usually, some of the synthesized sequences show high binding ability and specificity to target cells; some of them show binding ability to both target and control cell types; and some of them do not show binding affinity to both cell types. The sequences with high affinity and specificity are identified as aptamers for further characterization.

2.5.3 Characterization of Aptamers

Full-length aptamers generated by SELEX contain 60–120 nucleotides, which include two flanked primer sequences on each end for PCR amplification. Generally, not all the nucleotides are necessary for direct binding to target or for formation and maintenance of binding structure [31, 32]. Longer sequence would result in lower yield and higher cost in chemical synthesis of aptamers. The unnecessary nucleotides would also lead to higher chances of forming various secondary structures, thus destabilizing the target-binding conformation of aptamers. Thus, in practical usage, many selected aptamers are truncated down to a minimal functional sequence after the SELEX process. In general, the truncated sequences possess the same or better binding affinity than the original full-length aptamer. However, the problem here is how to truncate the full-length aptamer sequences or how to determine which nucleotides is unnecessary. RNase footprinting or partial hydrolysis has been used to determine the boundary and binding site of RNA aptamers [33–35]. Partially, fragmenting a full-length aptamer and then selecting the high affinity fragments have been used to determine the minimal sequence of DNA aptamers [36]. In these methods, radioactive labeling has been used to detect the aptamer fragments, and then the predicted potential minimal sequences had to be synthesized to confirm binding capacity.

In our cases, we utilize a relatively easy method to predict the secondary structure of aptamers, as well as the critical sequences for target binding [31, 32]. Secondary structure prediction of the aptamer sequences can provide information about relevant structures for binding [31, 32]. Such prediction is usually carried out by the program mfold [37] (http://www.idtdna.com/unafold/Home/Index), which calculates the possible secondary structures of single-stranded nucleic acids by energy minimizing method. The secondary structure prediction of a sequence often provides many potential structures; a few nucleotides or even single nucleotide change in a sequences can always leads to provide many different predicted structures. The sequences in same family should bind to same target with same secondary structures but different affinity. Therefore, comparing the predicted
structures of sequences in the same family, and even different families, can help to
determine the secondary structure for aptamer binding and the binding motif of
aptamers [31, 32]. Then, the aptamer sequences can be truncated based on their
predicted binding secondary structure and binding motifs. The truncated sequences
are synthesized and their binding ability is measured and compared with the origi-
nal full-length aptamers. The sequence truncation of aptamer may be carried out
many times to find out the minimized sequence. Based on the secondary structure
and binding motifs, the minimized aptamer sequences can be further optimized to
obtain better binding affinity by removing, adding, and/or displacing one or several
nucleotides [31, 32, 38–41].

When an optimized aptamer is obtained, its properties need to be further char-
acterized. The properties include dissociation constant ($K_d$), specificity to many
other cell lines, binding sites, tertiary binding structure, molecular target on cells,
and interactions with cells (such as cellular internalization [42]), as well as bio-
logical activities against cells. The $K_d$ values of an aptamer are usually measured by
flow cytometry [9]. It is important to test the $K_d$ values of aptamers at different
temperatures (such as 4, 25, and 37 °C), which would help the subsequent appli-
cations of aptamers. The details of cell-specific aptamer characterization and apt-
amer target identification are discussed in Chaps. 4 and 12.

2.6 Challenges and Perspectives

As discussed above, through cell-SELEX, aptamers that are able to recognize
molecular differences on cell membranes can be generated without prior knowledge
of their target molecules. These aptamers can be exploited for cell detection, cap-
ture, imaging, drug delivery, and biomarker discovery [43–46]. Since its devel-
opment, cell-SELEX has become an emerging and promising platform for
generating large numbers of aptamers against a wide range of cell lines, especially
various types of cancer cells. However, although considerable progress has been
made, cell-SELEX is still facing many challenges.

First, there is a big gap between theory and practice of cell-SELEX. The whole
SELEX process is very complex and like a “black box”; and no one knows exactly
at which step the problem occurred when no significant enrichment is observed
after many rounds of selection. The sequence discrimination in PCR amplification
of the random pools of DNA may greatly hinder the enrichment of aptamer
sequences. Additionally, some phenomena are unable to be explained currently; for
example, sometimes the high abundant sequences found in the final selected pool
do not bind to both target and control cell types. Actually, the success rate of cell-
SELEX is not high enough.

Second, the complexity and diversity of the cell surface ligands lead to the
difficulties in successful cell-SELEX attempts. Although in theory the selection of
aptamers is possible for any target on cell surface, there are preferences for some
types of target molecules to successfully select aptamers with high affinity and
specify. Target molecules with positive charges, hydrogen bond donors/acceptors, and/or aromatic groups facilitate aptamer selection. The high-abundance target molecules facilitate the aptamer selection. However, if these molecules on cell surface are not the desired targets (e.g., they are common molecules on many types of cells), they would disturb the enrichment of desired aptamers. Target molecules with negative charges or highly hydrophobic target molecules do not favor the aptamer selection. The low-abundance target molecules do not favor the aptamer selection. The abundance of some membrane molecules may change with the cell growth status and cell passages, which would interfere in the aptamer selection. Therefore, the good cell culture maintenance is important for cell-SELEX.

Third, in contrast to the traditional SELEX against purified targets, a target identification step is necessary for cell-SELEX after obtained desired aptamers. However, until now, only a limited number of aptamer targets have been identified [11, 47–49], which greatly hindered the further application of aptamers selected against cells. The separation and purification of aptamer targets, especially the membrane protein targets are still a serious limitation for the target identification.

Fourth, the long selection period, complex and tedious process, high experience requirement, and high cost of SELEX still restrict the widely utilization of cell-SELEX.

As a result, some modified cell-SELEX strategies have been reported to exhibit improved selection efficiency. FACS-SELEX integrating flow cytometry and cell sorting into cell-SELEX allows simultaneously removing selection-hampering dead cells and selecting aptamers that are able to differentiate cell subpopulations [50, 51]. An on-chip Cell-SELEX process for automatic selection of aptamers has shorten the selection process to five rounds [52]. In the future for a long period of time, continued efforts on cell-SELEX need focus on the following aspects: (1) Reveal the principle of aptamer enrichment in cell-SELEX; reveal the binding mechanisms of aptamers and the targets on cell surface. (2) Develop new methods to enhance the selection efficiency, shorten the selection period and lower the cost, such as higher automated selection platform. (3) Develop efficient and universal strategies for the target identification of aptamer. (4) Develop data library of aptamers that generated by cell-SELEX, which include the information of target and control cell lines, initial library, selection condition, aptamer sequences, binding affinity, and specificity. This library will provide comprehensive information for further understanding of cell-SELEX and cellular biology.

Although the cell-SELEX methods need to be further improved, current researches have demonstrated that cell-SELEX can do a lot of interesting things, which is encouraging more researchers to develop additional aptamers for different targets. Despite compared with the complex molecular changes on cancer cell surface, the aptamers obtained by cell-SELEX are too less, we believe that further development of cell-SELEX and accumulation of aptamers will lead to an improved understanding of the biochemical and molecular basis of cancer, in turn, spur exciting new technologies for detection, diagnosis, and treatment of cancer. The cell-SELEX strategy can also be applied to the investigation of other diseases, thus expanding our knowledge of disease biology and improving the medical care and
life expectancy of individual patients. We are sure that along with the development of cell-SELEX, it will play a more important role in many fields such as clinical detection, personalized therapy, and cellular biological research.

References


Aptamers Selected by Cell-SELEX for Theranostics
Tan, W.; Fang, X. (Eds.)
2015, XIII, 352 p. 116 illus., 107 illus. in color.,
Hardcover
ISBN: 978-3-662-46225-6