Published in final edited form as:

Analyst. 2016 January 21; 141(2): 461–466. doi:10.1039/c5an01918d.

Cancer Biomarker Discovery Using DNA Aptamers

Cheng Jin^{a,%}, Liping Qiu^{a,b,%,†}, Jin Li^a, Ting Fu^{a,b}, Xiaobing Zhang^a, and Weihong Tan^{a,b,†}

^a Molecular Science and Biomedicine Laboratory, State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, College of Biology, Collaborative Innovation Center for Molecular Engineering for Theranostics, Hunan University, Changsha, 410082, China

^b Center for Research at Bio/nano Interface, Department of Chemistry and Department of Physiology and Functional Genomics, Health Cancer Center, UF Genetics Institute and McKnight Brain Institute, University of Florida, Gainesville, FL 32611-7200, United States

Abstract

Biomarkers are signature molecules able to indicate specific physiological states of cells. Identification of reliable biomarkers is essential for early diagnosis and adaptive treatment of diseases, especially cancer. Aptamers are single-stranded oligonucleotides generated by an *in vitro* screening method called Systematic Evolution of Ligands by Exponential Enrichment (SELEX). They can recognize their cognate targets with selectivity and affinity comparable to that of protein antibodies. In addition, aptamers have superiorities including easy synthesis, high chemical stability, convenient modification and flexible design. As such, these DNA molecules show great promise as powerful molecular probes for biomarker discovery and biomarker-based clinical applications. Using complex samples as targets, a panel of aptamers can be systematically generated for comprehensive recognition of disease-specific proteins, which can potentially serve as biomarkers. This review describes current methods for biomarker discovery using aptamers.

INTRODUCTION

Biomarkers, as specific indicators of cellular states, have become increasingly important in understanding the molecular basis of diseases, cancer in particular, and in developing early diagnosis and adaptive therapies.^{1, 2} Despite intense efforts from academia and industry, as well as significant financial investment, very few biomarkers have met the requirements for routine clinical use, even for the most common cancers. In the past decade, the only tumor biomarker approved by the U.S. Food and Drug Administration (FDA) is the prostate-specific antigen (PSA) for prostate cancer.³ This slow progress can be attributed to the complexity and heterogeneity of cancers, as well as the lack of efficient methods for biomarker discovery.

[†]To whom correspondence should be addressed: tan@chem.ufl.edu qiuliping@hnu.edu.cn Tel/Fax: +1 352 846 2410.
[%]equal first author contribution

Electronic Supplementary Information (ESI) available: See DOI:10.1039/x0xx00000x

The authors declare no competing financial interest.

Conventional strategies for biomarker discovery primarily involve genomics and proteomics.^{4,5} Although investigations of gene expression can be informative, they do not always correlate with the abundance and the variation of encoded proteins, which are often subject to proteolytic cleavage or post-translational modification. Proteomic approaches, which focus on characterizing the expression and alteration of proteins, are more straightforward for interpreting the cellular phenotype.⁶ By combining mass spectrometry (MS) with two-dimensional gel electrophoresis (2D-GE), some progress has been made in the proteome analysis and the specific protein identification. However, a common limitation still remains in that these technologies tend to provide false positive results originating from intensive sample processing, including cell/tissue lysis, isolation and purification.^{7,8} Thus, even small operational differences during sample processing can lead to large deviations in the results. In addition, while MS can address specific analytical capabilities, current MS-based techniques have such shortcomings as low sensitivity with LODs, typically in the nanomolar level, and low specificity, making it challenging to identify low-abundance biomarkers in protein complex.⁹ Moreover, the downstream applications of biomarkers, such as tumor imaging and targeted drug delivery, require the development of additional molecular probes, further impeding implementation of biomarker-based assays in clinical medicine. In view of these difficulties, a variety of cancer-specific molecular probes is needed for efficient biomarker discovery and practical applications.

Antibody-antigen interactions have the high binding affinities with dissociate constant (K_d) that range from nanomolar to picomolar. As such, the incorporation of antibodies as recognition elements is expected to increase assay sensitivity to subnanomolar level.¹⁰ However, antibody-based strategies often require a sandwich system for screening target proteins. Thus, to identify one target protein, two different antibodies are required, making antibody-based systems impractical for large-scale biomarker discovery. Moreover, systematic production of antibody libraries for molecular differentiation of cancer cells is very difficult. For most diseases, few reliable molecular probes are specific enough to recognize subtle molecular differences among closely related diseases.¹¹

Aptamers, as chemical antibodies, are emerging as important tools in molecular recognition. Aptamers, which are artificial DNA/RNA strands selected through an *in vitro* method named SELEX (Systematic Evolution of Ligands by Exponential Enrichment), can form into specific tertiary structures to selectively recognize various targets ranging from small molecules to peptides and proteins and even whole cells.¹² Apart from their comparable affinity and specificity with protein antibodies, aptamers have additional advantages including easy synthesis, high batch-to-batch reproducibility, low cost, convenient modification with various functional groups, flexible design and low immunogenicity, making them promising molecular probes for biological applications.

APTAMER-BASED BIOMARKER DISCOVERY

Using purified proteins as targets, several aptamers have been selected against disease-related proteins such as HER3, thrombin, PDGF, VEGF, PMSA, tenascin-C and NF κ B. While these aptamers have been widely used in bioanalysis and biomedicine, aptamer selection against complex targets, such as serum, whole cells and tissue slices, is medically

more relevant.^{13, 14} Gold *et al* developed a DNA aptamer-based proteomics platform to screen thousands of proteins in small-volume biological samples, either serum or plasma.¹⁵ They first synthesized a random oligonucleotide library with chemically-modified nucleotides to increase chemical diversity and then developed the corresponding SELEX method to select aptamers with slow off-rates. They developed a multiplexed system by using these modified oligonucleotides to conduct multi-rounds selection against a complex protein sample, and demonstrated that the selected aptamers were able to measure 813 proteins with low limits of detection (1 pM) and wide dynamic ranges (over 7 orders of magnitude). Furthermore, they used the selected aptamers to set up a proteomics platform to screen biomarkers for chronic kidney disease (CKD). Two previously recognized CKD biomarkers, as well as 58 potential CKD biomarkers, were identified, demonstrating the potential of this platform for biomarker discovery on a large scale. In another study, the Gold group used the same technique for biomarker discovery for non-small cell lung cancer among a group of long-time smokers, and 44 potential serum biomarkers of lung cancer were identified.¹⁶ With the information provided by these biomarkers, they developed an aptamer-based proteomics array which enabled the selective discrimination of cancer samples from normal cells.

CELL-SELEX-BASED BIOMARKER DISCOVERY

While serum biomarkers are beneficial for simplifying cancer diagnosis in the form of blood tests, cellular markers can be indispensable for elucidating cell biology, *in vivo* tumor imaging and targeted anticancer drug delivery. Recently, a new application of the SELEX technique termed Cell-SELEX was reported by Tan *et al.*, using whole cells as the targets to explore the different patterns of cell membrane epitopes for two given cell lines.^{17, 18} Within the cell-SELEX process, cell-surface proteins remain in their native conformations, which is essential for retaining their intact biological functions. Cell-based SELEX can also be used to select cell type-specific aptamers able to recognize the molecular signatures of cells with no prior knowledge of the target receptors. Moreover, by targeting an entire cell membrane having various receptors, a panel of aptamers can be systematically generated to profile characteristics of the target cells, resulting in a higher dimension of information content for cancer biomarker discovery.

A general process of cell-based SELEX is illustrated in Figure 1.¹⁹ First, the target cells are incubated with an oligonucleotide library of over 10^{15} sequences. After removal of unbound sequences with several rounds of washing and centrifugation, the bound probes are eluted by heat treatment. Then excess control cells are added to interact with the eluted DNA pool to eliminate the sequences which bind to molecules common to both cell lines, ensuring that the pool contains only sequences that specifically bind to the target cells but not the control cells. The collected sequences are then amplified by PCR to generate a new DNA pool for the next round of selection. Selection progress can be evaluated by performing binding assays using flow cytometry. Generally, after nearly 20 rounds of selection, some specific DNA probes are enriched, and the corresponding pool is then sequenced.

Using the Cell-SELEX technology, Tan *et al.* have selected a group of aptamers that can specifically recognize CCRF-CEM cells, a human T-cell acute lymphoblastic leukemia cell

line, with K_d s in the nanomolar to picomolar range.¹⁷ They demonstrated the ability of these aptamers to differentiate between the target leukemia cells and other cancer cell lines (e.g., Ramos, a Burkitt's lymphoma cell line), as well as a mixture of normal human bone marrow aspirates. Subsequently, they used one of the representative aptamers, sgc8 ($K_d \approx 0.8$ nM), as the capture probe to identify its target receptor from the CCRF-CEM cell lysate. The sgc8 aptamer was labeled with a biotin tag and used to extract its receptor with the aid of streptavidin-coated magnetic beads (as shown in Figure 2). The collected receptors were then analyzed by LC-MS/MS, and protein tyrosine kinase 7 (PTK7), a transmembrane receptor, was identified. It has been reported that PTK7 is overexpressed in several tumors, although its exact role in tumor formation is not well documented.²⁰ Aptamer sgc8 provides a versatile tool to investigate the biological function of PTK7. For example, with programmability and easy modification, sgc8 can be used for real-time monitoring of the cellular distribution of PTK7, cell-targeted delivery of nano/molecular therapeutic agents, cancer profiling, and manipulation of biological activity of PTK7 through combination with inhibition modules.

In another work, the Tan group used a similar strategy to identify the target receptor of the TDO5 aptamer which was selected by using Ramos cells as the target.²¹ To ensure capture efficiency, TDO5 was modified with a photocrosslinking moiety. After specific binding, TDO5 was covalently crosslinked with its receptor under light activation. With this covalent crosslinking reaction, the binding affinity of TDO5 with its receptor can be greatly enhanced, avoiding possible off-target events during the protein extraction process. They verified that the target of TDO5 was immunoglobulin heavy mu chain (IGHM), which was previously reported to be closely associated with the development of Burkitt's lymphoma (as shown in Figure 3).²² Both of these studies are considered promising steps in cell-type specific aptamer-based biomarker discovery.

Acute myelogenous leukemia (AML) is one of the most challenging cancers for human being with rather low survival rate. To discover biomarkers for the diagnosis and therapy of AML, Li's group applied the Cell-SELEX technique on the NB4 AML cell and generated a DNA aptamers pool for the molecular recognition of this cell line.²³ Three aptamers that can selectively recognize the NB4 AML cell with K_d in the low nanomolar range were characterized. They used these aptamers to phenotype AML cells in clinical specimens, and identify a low-abundance biomarker, Siglec-5, for AML.

Dendritic cells (DCs) are antigen-presenting cells that function to activate lymphocytes of the immune system.²⁴ They have become one of the research focus in generating therapeutic immunity against cancers. Analyzing the molecular features of DCs in specific biological states can provide useful information in designing DCs-based immunotherapeutic strategies for cancer treatments. Initially, DCs are in an immature state, but when stimulated by infection or cancers, immature DCs (iDCs) grow into mature DCs (mDCs) by regulation of their phenotypes to initiate the adaptive immune response of lymphocytes. It follows that biomarkers able to distinguish DCs in their different states would be critical for understanding the biological mechanisms of the immune system. Accordingly, Krylov *et al.* developed an aptamer-based strategy to identify biomarkers for both iDCs and mDCs. They first used mDCs as the positive cell line and iDCs as the negative cell line to select a mDC-

specific aptamer pool. Then they used the selected aptamer pool to extract biomarkers from mDCs. By analyzing the collected proteins by LC-MS, they identified two well-known biomarkers and one potential candidate for mDCs. Meanwhile, using a similar strategy to isolate an aptamer pool specific for iDCs, using iDCs as the target and mDCs as the control, the authors identified six potential biomarkers for iDCs. In addition, they found that a pool of aptamers was more efficient than an individual aptamer for both cell isolation and biomarker discovery.

Selection of Aptamers for Known Biomarkers

The development of high-performance aptamers which can recognize and signal the identified biomarkers is important for the practical applications of these biomarkers. Epithelial cell adhesion molecule (EpCAM) is mostly overexpressed on the membrane of solid tumor cells and is applied as a significant biomarker in clinical cancer diagnosis and treatments. By using human recombinant protein as the target, Yang's group has successfully selected a panel of DNA aptamers against EpCAM.²⁵ They demonstrated that these aptamers can specifically recognize the EpCAM-positive cancer cells but not bind to the EpCAM-negative cells (as shown in Figure 4). They also used the optimized aptamers SYL3C to capture the target cancer cells from the cell mixture and achieved 63% target cell capture efficiency and 80% cell purity. Such EpCAM-specific aptamers show great potential for cancer cell imaging, target drug delivery and target cell capture.

Prostate specific antigen (PSA) is a FDA-approved biomarker for prostate cancer, and is highly valuable as the reference to the diagnosis of such cancer. By using the traditional SELEX technique, Ikebukuro's group pre-screened a PSA-specific aptamer pool.²⁶ To improve the binding affinity of the aptamers, they introduced a genetic algorithm (GA) to the post-SELEX screening. GA is an efficient algorithm which generates solutions to optimization problems using techniques inspired by natural evolution. GA can be used to improve any features of aptamers, for example, binding affinity and selectivity. With introduction of GA, they obtained a PSA-specific aptamer with a dissociation constant in the tens of nanomolar range, and successfully analyzed PSA at the concentration from 40 to 100 nM. On the other hand, the PSA test is controversial in clinical cancer diagnosis. The aptamer system, which uses the prostate cancer sample as the selection target and the relative normal sample as a control, would provide more molecular features of prostate cancer in addition to PSA, thus potentially overcoming the clinical shortcomings of the current PSA test.

Summary and Perspectives

Based on the SELEX technology, a panel of aptamers can be systematically generated for complex biological systems, such as serum, cells and tissue, thus enabling the comprehensive recognition of molecular signatures of specific diseased samples. Theoretically, aptamers can discriminate subtle differences between the target sample and the control, thus, to some extent, avoiding false positive results. In addition, taking advantage of flexible design and convenient modification, aptamers can be easily applied for practical biomarker-based assays, making the identified biomarkers applicable in clinical settings. Moreover, with combination of high-throughput nucleic acid-based devices, such as

DNA microarrays and DNA nanochips, it is possible to obtain aptamer-based biomarker profiling on individual cancer patient samples, thus overcoming, to some extent, issues related to the heterogeneity of cancer. In this light, aptamers may bring a revolutionary change in personalized cancer medicine.

Despite great progress, aptamer-based cancer biomarker discovery is still in the infant stage. For further development of this field, some issues should be addressed. First, a larger number of aptamers specific for cancer-related samples are needed to facilitate the discovery of more cancer biomarkers. Second, methods to enhance the binding affinity of aptamers with their targets can effectively reduce off-target events during protein extraction from the complex sample (e.g., cell lysate). In addition, since nucleic acids tend to be digested by various nucleases in biological systems, chemical modification, which can improve the stability of aptamers with no negative impact on their selectivity and binding affinity to the targets, are essential for the practical application of aptamers in cancer diagnosis and therapies. Moreover, the aptamer selection process always takes more than one month, and the results show some degree of randomness, hampering the development of aptamer-based biomarker discovery to some extent. In this light, improvement on the SELEX techniques is urgently needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work is supported by the National Key Scientific Program of China (2011CB911000), NSFC grants (NSFC 21221003 and NSFC 21327009) and China National Instrumentation Program 2011YQ03012412, and by the National Institutes of Health (GM079359 and CA133086).

Notes and references

- Henriksen K, O'Bryant SE, Hamper H, Trojanowski JQ, Montine TJ, Jeromin A, Blennow K, Lonneborg A, Wyss-Coray T, Soares H, Bazenet C, Sjogren M, Hu W, Lovestone S, Karsdal MA, Weiner MW, G. Blood-Based Biomarker Interest. *Alzheimers & Dementia*. 2014; 10:115–131.
- Frangogiannis NG. *Translational Research*. 2012; 159:197–204. [PubMed: 22424424]
- Fuzery AK, Levin J, Chan MM, Chan DW. *Clinical proteomics*. 2013; 10:13–13. [PubMed: 24088261]
- Liu R, Chen X, Du Y, Yao W, Shen L, Wang C, Hu Z, Zhuang R, Ning G, Zhang C, Yuan Y, Li Z, Zen K, Ba Y, Zhang C-Y. *Clinical Chemistry*. 2012; 58:610–618. [PubMed: 22194634]
- Whiteaker JR, Lin C, Kennedy J, Hou L, Trute M, Sokal I, Yan P, Schoenherr RM, Zhao L, Voytovich UJ, Kelly-Spratt KS, Krasnoselsky A, Gafken PR, Hogan JM, Jones LA, Wang P, Amon L, Chodosh LA, Nelson PS, McIntosh MW, Kemp CJ, Paulovich AG. *Nature Biotechnology*. 2011; 29:625–U108.
- Skates SJ, Gillette MA, LaBaer J, Carr SA, Anderson L, Liebler DC, Ransohoff D, Rifai N, Kondratovich M, Tezak Z, Mansfield E, Oberg AL, Wright I, Barnes G, Gail M, Mesri M, Kinsinger CR, Rodriguez H, Boja ES. *Journal of Proteome Research*. 2013; 12:5383–5394. [PubMed: 24063748]
- Wasinger VC, Zeng M, Yau Y. *International journal of proteomics*. 2013; 2013:180605–180605. [PubMed: 23533757]
- Indovina P, Marcelli E, Pentimalli F, Tanganelli P, Tarro G, Giordano A. *Mass Spectrometry Reviews*. 2013; 32:129–142. [PubMed: 22829143]

9. Drucker E, Krapfenbauer K. The EPMA journal. 2013; 4:7–7. [PubMed: 23442211]
10. Baleviciute I, Balevicius Z, Makaraviciute A, Ramanaviciene A, Ramanavicius A. Biosensors & Bioelectronics. 2013; 39:170–176. [PubMed: 22938839]
11. de Castro DG, Clarke PA, Al-Lazikani B, Workman P. Clinical Pharmacology & Therapeutics. 2013; 93:252–259. [PubMed: 23361103]
12. Wilson DS, Szostak JW. Annual review of biochemistry. 1999; 68:611–647.
13. Shi H, Cui W, He X, Guo Q, Wang K, Ye X, Tang J. Plos One. 2013; 8
14. Shamah SM, Healy JM, Cload ST. Accounts of Chemical Research. 2008; 41:130–138. [PubMed: 18193823]
15. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney TR, Kim N, Koch TH, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson SK, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff RM, Otis M, Parker T, Pietrasiewicz S, Resnicow DI, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught JD, Vrkljan M, Walker JJ, Watrobka M, Waugh S, Weiss A, Wilcox SK, Wolfson A, Wolk SK, Zhang C, Zichi D. Plos One. 2010; 5
16. Lambris JD, Holers VM, Ricklin D, Ricklin D. Lambris JD, Holers VM. Complement Therapeutics. 2013; 735:V-V.
17. Shangguan D, Li Y, Tang Z, Cao ZC, Chen HW, Mallikaratchy P, Sefah K, Yang CJ, Tan W. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103:11838–11843. [PubMed: 16873550]
18. Fang X, Tan W. Accounts of Chemical Research. 2010; 43:48–57. [PubMed: 19751057]
19. Sefah K, Shangguan D, Xiong X, O'Donoghue MB, Tan W. Nature Protocols. 2010; 5:1169–1185. [PubMed: 20539292]
20. Shangguan D, Cao Z, Meng L, Mallikaratchy P, Sefah K, Wang H, Li Y, Tan W. Journal of Proteome Research. 2008; 7:2133–2139. [PubMed: 18363322]
21. Tang Z, Shangguan D, Wang K, Shi H, Sefah K, Mallikaratchy P, Chen HW, Li Y, Tan W. Analytical Chemistry. 2007; 79:4900–4907. [PubMed: 17530817]
22. Mallikaratchy P, Tang Z, Kwame S, Meng L, Shangguan D, Tan W. Molecular & Cellular Proteomics. 2007; 6:2230–2238. [PubMed: 17875608]
23. Yang M, Jiang G, Li W, Qiu K, Zhang M, Carter CM, Al-Quran SZ, Li Y. Journal of Hematology & Oncology. 2014; 7
24. Berezovski MV, Lechmann M, Musheev MU, Mak TW, Krylov SN. Journal of the American Chemical Society. 2008; 130:9137–9143. [PubMed: 18558676]
25. Song Y, Zhu Z, An Y, Zhang W, Zhang H, Liu D, Yu C, Duan W, Yang CJ. Analytical Chemistry. 2013; 85:4141–4149. [PubMed: 23480100]
26. Savory N, Abe K, Sode K, Ikebukuro K. Biosensors & Bioelectronics. 2010; 26:1386–1391. [PubMed: 20692149]

Figure 1. Schematic representation of DNA aptamer selection using the cell-SELEX strategy. Reproduction reprint permission of *Nat. Protoc.* copyright 2010.¹⁹

Figure 2.

PAGE analysis of sgc8c-assisted target purification. Lane 1, DNA markers; lane 2, cell membrane lysate; lane 3, protein captured with the random DNA; lane 4, magnetic beads only; lane 5, protein captured with DNA sgc3b; lane 6, protein captured with aptamer sgc8c. Reproduction reprint permission of *J. Proteome Res.* copyright 2008.²⁰

Figure 3.
 Schematic representation of aptamer-based identification of IGHM on Ramos cells.
 Reproduction reprint permission of *Mol. Cell. Proteomics* copyright 2007.²²

Figure 4. Selection of DNA aptamers against EpCAM and their application in cancer cell imaging and Circulating Tumor Cell (CTC) capture. Reproduction reprint permission of *Anal. Chem.* copyright 2013.²⁵