

In silico selection approach to develop DNA aptamers for a stem-like cell subpopulation of non-small lung cancer adenocarcinoma cell line A549

Mateja Vidic^{1,2*}, Tina Smuc^{3**}, Nika Janez³, Michael Blank⁴, Tomaz Accetto⁵, Jan Mavri^{3**}, Isis C. Nascimento⁶, Arthur A. Nery⁶, Henning Ulrich⁶, Tamara T. Lah^{1,7}

¹ Department of Genetic Toxicology and Cancer Biology, National Institute of Biology, Ljubljana, Slovenia

² Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

³ Centre of Excellence for Biosensors, Instrumentation and Process Control, Centre for Biotechnology, Ajdovščina, Slovenia

⁴ AptalT GmbH, Munich, Germany

⁵ Department of Animal sciences, Biotechnical Faculty, University of Ljubljana, Slovenia

⁶ Department of Biochemistry, Institute of Chemistry, University of São Paulo, Brazil

⁷ Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

* Current address: Aptamer Group, York YO10 5NY, United Kingdom

** Current address: Lek-Sandoz Company, Ljubljana, Slovenia

Radiol Oncol 2018; 52(2): 152-159.

Received 24 November 2017

Accepted 26 February 2018

Correspondence to: Prof. Henning Ulrich, Department of Biochemistry, Institute of Chemistry, University of São Paulo, Av. Prof. Lineu Prestes 748, São Paulo, 05508-000 SP, Brazil. E-mail: henning@iq.usp.br and Prof. Tamara T. Lah, National Institute of Biology, Večna pot 111, 1000 Ljubljana, Slovenia. E-mail: Tamara.lah@nib.si

Disclosure: No potential conflicts of interest were disclosed.

Background. Detection of circulating lung cancer cells with cancer-stem like characteristics would represent an improved tool for disease prognosis. However, current antibodies based methods have some disadvantages and therefore cell SELEX (Systematic Evolution of Ligands by Exponential Enrichment) was used to develop DNA aptamers, recognizing cell surface markers of non-small lung carcinoma (NSLC) cells.

Materials and methods. The human adenocarcinoma cell line A549 was used for selection in seven cell SELEX cycles. We used human blood leukocytes for negative selection, and lung stem cell protein marker CD90 antibody binding A549 cells for positive selection.

Results. The obtained oligonucleotide sequences after the seventh SELEX cycle were subjected to *in silico* selection analysis based on three independent types of bioinformatics approaches, selecting two closely related aptamer candidates in terms of consensus sequences, structural motifs, binding affinity (Kd) and stability (ΔG). We selected and identified the aptamer A155_18 with very good binding characteristics to A459 cells, selected CD90 antibody binding. The calculated phylogenetic tree showed that aptamers A155_18 and the known A549 cell aptamer S6 have a close structural relationship. MEME sequence analysis showed that they share two unique motifs, not present in other sequences.

Conclusions. The novel aptamer A155_18 has strong binding affinity for A549 lung carcinoma cell line subpopulation that is expressing stem cell marker CD90, indicating a possible stemness, characteristic for the A459 line, or a subpopulation present within this cell line. This aptamer can be applied as diagnostic tool, identifying NSLC circulating cells.

Key words: aptamers; cell-SELEX; circulating tumour cells; non-small lung cancer; computational biology

Introduction

DNA aptamers acquire oligonucleotides' tertiary structures that allow them to bind to various target macromolecules, such as proteins, via non-covalent binding. Aptamer binding affinity to cell surface epitopes is in the range of monoclonal antibodies.^{1,2} When designed to bind to the whole cell, it is assumed that aptamers bind either to specific proteins or other complex molecular structures. The tight binding aptamers are selected by an *in vitro* selection method in a stepwise process from a starting combinatorial library of 10¹⁵ random oligonucleotides by a process called SELEX (Systematic Evolution of Ligands by Exponential Enrichment), reviewed by Šmuc and Ulrich.³ In brief, the SELEX process comprises five main steps: binding, partition, elution, amplification, and conditioning in a reiterative and stepwise manner, which is narrowed down to a homogeneous population of high-target affinity and selectivity sequences.⁴ The here used Cell-SELEX approach is a modification of the original of the original method, where a selection of aptamers binding to cell surface epitopes of target cells, i. e. tumor cells, is followed by a negative selection step against a non-target cells in order to remove any sequences, which commonly bind to epitopes, expressed by both cell types. Finally obtained aptamers shall be selective for the desired cell type.^{5,6}

Although basic mechanisms of aptamer target binding are known, theoretic prediction of individual oligonucleotide binding to cellular surfaces cannot be done. However, novel bioinformatics tools have been developed recently that can discriminate among an already selected set of aptamers with the lowest dissociation constant and the highest binding energy. In cancer, aptamers have been suggested to replace the antibodies mostly for diagnostic purposes, as they are more reliable in terms of reproducibility, stability, and costs of production⁷.

The detection of circulating tumour cells (CTC) in body fluids prior to, or at the first medical intervention, would represent a particular challenge for the prediction of disease progression⁸. The two technologies used for CTC enumeration, the Cell Search System[®] based on the detection of cancer cell membrane protein markers by antibodies⁹ and the platform ISET (Isolation by Size of Epithelial Tumour cells), based on cell size exclusion, are neither by themselves, nor used complementarily sufficient for CTC based diagnosis⁹, therefore new approaches are needed. Lung cancer incidence and

death rates are still increasing. The subgroup of NSCLC appears to have the highest incidence rates and is mostly locally advanced or metastatic at the time of diagnosis.¹⁰ Therefore we have used the cell line A549, established from the primary tumour of a NSCLC patient, to raise the aptamers. There have been several successful attempts so far to target NSCL cells in the blood circulation (lung CTC) by specific aptamers.¹¹⁻¹³ However, these did not address the potential stemness of CTC, which appeared to discriminate among cells with the highest tumorigenic potential and is thus more relevant for aggressive progression and worse prognosis of lung cancer. These lung cancer stem cells (CSC) express high levels of CD44^{high} and CD90⁺ protein.¹⁴ Furthermore, it has been shown that CD90⁺ A549 cells also express CSC markers, such as Oct4, Sox2 and some others.¹⁴ These cells had higher proliferation rates and tumorigenic capacities, and Yan *et al.* assumed that among lung cancer patients a subpopulation of lung CSC cells exists, which could be detected by specific aptamers. The aim of this study was thus the development of DNA aptamers that would also bind cancer-stem like cell surface biomarkers and would be suitable to detect stem cell-like CTC in blood circulation. The novelty in this study is a more efficient development, of aptamer binding to target cell surface based also on a *in silico* selection by novel bioinformatics tools used for the first time, and suggests their application in future aptamer identifications.

Materials and methods:

Cell culture and reagents

Human lung carcinoma cell line A549 (ATCC[®] CCL-185TM, ATCC, Manassas, VA) at passage 60 was cultured in DMEM (Millipore Sigma, Burlington, USA) supplemented with 10% foetal bovine serum until they reached about 80% confluence. Cells were washed to remove residuals of medium and then detached from the bottles with 2mM EDTA solution (Millipore Sigma, Burlington, MA). The cell line authentication was performed with IdentiCell STR allele protocol and showed a 100% match with A549 cells (IdentiCell, Department of Molecular Medicine, Aarhus, Denmark).

Cell-SELEX library design

A random library (5'-FITC- GCC TGT TGT GAG CCT CCT-N34-CGC TTA TTC TTG TCT CCC-3')

regions for the binding of primers during the PCR amplification reactions was used.¹⁵ The forward primer was 5'-BBB-GCC TGT TGT GAG CCT CCT-3', where BBB indicates 3 subsequent biotin moieties, while the reverse primer was labelled with a 6-FAM or Fluorescein Isothiocyanate (FITC-5'-GGG AGA CAA GAA TAA GCG-3'). Before the first selection cycle, the library was amplified by a PCR reaction (PCR Conditions described in Supplementary material), subjected to denaturation and the fluorescence single-stranded DNA was purified by denaturation PAGE. At the beginning of the SELEX cycle, the aptamer pool in selection buffer was denatured at 95°C for 10 min and then placed immediately on ice for 10 min followed by 20 min incubation at room temperature.

Selection of aptamers

The incubation of cells with the aptamers was followed by the BRAZIL technique with a centrifugation of a mixture of cells and oligonucleotides through a dibutyl phthalate: cyclohexane (9:1 [v:v]; d=1.03 g.ml⁻¹) layer¹⁶.

Briefly, the cells with bound aptamers were collected from the pellet of the organic phase and separated by centrifugation (13,000 g for 10 min (Hettich Universal 32R centrifuge, HETTICH Instruments LP, Beverly, MA). The pellet containing aptamers was extracted by phenol/chloroform (1:1) for purification of obtained DNA followed by PCR as detailed above and strand separation for the next cycle of SELEX. To perform selection of ligands with higher affinity, the stringency of the selection was gradually increased in each cycle by adding 0.3 – 3 mg/ml tRNA for reducing nonspecific binding and increasing the ratio of DNA molecules over cells (10⁶ – 10⁵ cells).

In the seventh cycle the first negative selection with human blood cells (erythrocytes, leukocytes, and thrombocytes in the amount of 10⁷ cells) was performed, followed by a positive selection step against 10⁵ A549 cells. Blood cells were separated from plasma by density gradient separation.¹⁷ A selection step against a CD90⁺ A549 cells was done following cell sorting purification of this subpopulation.

Flow cytometry analysis of aptamer binding to A549 cells

These assays were performed with live A549 cells, which not been treated with fixation agents according to Nascimento *et al.* (2016).¹⁸ Pools were incu-

bated in binding buffer (1.25mM HEPES, 0.27mM KCl, 0.14mM CaCl₂, 0.06mM MgCl₂, 7.19mM NaCl, 0.9% glucose) for 10 min at 95°C and then immediately placed on ice for 10 min and maintained for 20 min at room temperature. For the interaction between single-stranded DNA molecules and cells, the solution containing the aptamers was gently mixed with cells (10⁶ cells /microtube) to a final volume of 100 µl and 200 pM concentration of aptamers. The mixture was incubated for 30 min at room temperature with gentle shaking. The mixture was centrifuged at 200 X g and the supernatant was discarded. A wash was performed with binding buffer, the mixture was again centrifuged and the supernatant discarded. Then 500 µl of binding buffer was added, followed by flow cytometry analysis. For analysis of double-labelled cells, the anti-CD90 antibody (1: 1,000) was added to the cells along with the oligonucleotide pools. For cell sorting the same procedure was followed, and a gate was determined to collect the double-labelled cells for CD90 expression and aptamer oligonucleotides.

The pool from the seventh cycle contained aptamers with enhanced binding to the subpopulation of A549 cells, expressing CD90, as was detected by increased fluorescence signals. Since the identified subpopulation demonstrated stem-like cells characteristics¹³, cell sorting was used to isolate the identified fraction of cells. A fraction of 2,500 cells was successfully collected and aptamers were purified for PCR amplification (data not shown).

Sequencing

Selection pool aptamers obtained in each cycle were tested with restriction analysis (RFLP – Restriction Fragment Length Polymorphism) for the presence of conserved restriction sites. Different bands were obtained representing fragments of DNA with conserved restriction sites, thus showing selective enrichment of specific groups of oligonucleotides during SELEX (Suppl. Figure S1).

The sequences from last three cycles were used for Sanger and Next generation sequencing (Supplementary material). Selected sequences were further analysed with different bioinformatics approaches.

Bioinformatics analysis

Three different *in silico* selection procedures were used for processing the selected sequences: CLC Genomics Workbench software, Shell script in

combination with pipelined use programs for processing nucleotide sequences- MEME¹⁹, MEGA6²⁰ and UnaFold²¹ and COMPAS software (AptaIT GbmH, Planegg, Germany).²²

We have compared molecular evolutionary relations (with MEGA6 program²⁰) and motif similarity (with MEME suite¹⁹) of chosen four best binding candidates (A155_18, A452_3, A373_4, A218_12) and already published aptamer S6 binding to A549 cells (5'-GTG GCC AGT CAC TCA ATT GGG TGT AGG GGT GGG GAT TGT GGG TTG-3').¹¹

Validation of selected candidates *in vitro*

Preparation of aptamers: Aptamers were dissolved at 0.5 μ M final concentration in binding buffer, containing 25mM HEPES (Millipore Sigma) 5.4 mM KCl, 2.8 mM CaCl₂, 3.2 mM MgCl₂ and 144 mM NaCl in ddH₂O and denaturated at 95°C and renatured at room temperature, each for 20 min. Further, the cells were incubated for 20 min at 25°C with 0.5 μ M 6-Carboxyfluorescein (6-FAM) labelled aptamer solution in 200 μ l of binding buffer with gentle agitation. Then the cells were washed twice with binding buffer before flow cytometry analysis of A459 cell-aptamer bindings. For determination of unspecific binding, we used a 6'-FAM oligonucleotide with a random region of 34 nucleotides. To determine whether binding between aptamer and the target cells would depend on temperature, binding assays were also carried out at 4°C and 37°C.

CD90 antibody binding to A549 cells: To determine CD90 protein expression in A549 cells, they were incubated with the anti-CD90/Thy1 antibody (PE/Cy5) (ab95698) and for negative control assays with Isotype Control Mouse IgG1, Kappa monoclonal (PE/Cy5) (ab67435) obtained from AbCam (Cambridge, UK). The cells were detached by 0.02% EDTA and filtered through a 70 μ m sieve to avoid doublets. Cells – 300,000 per assay were washed three times with ice cold phosphate-buffered saline (PBS) and incubated for 30 min at 4°C and then re-suspended in 500 μ l ice cold PBS before flow cytometry analysis.

Preparation of blood cells for negative control: As the aim of aptamer developments was to detect NSCLC cells in the blood as CTC, the blood cells were isolated from a healthy donor (male, 40 years) lysed with buffer, containing 150 mM NH₄Cl, 10mM NaHCO₃ and 0.1mM EDTA and centrifuged at 300 \times g for 5 min at room temperature. The blood cells were re-suspended in binding buffer and used as binding target for the negative selection step with previous-

ly eluted SELEX pool DNA to separate aptamers with affinity to the tumour cell target and to discard sequences also binding to control blood cells. The remaining target cell specific sequences were further PCR amplified to form the starting pool for the final round of positive selection.

Flow cytometry: We measured 30,000 events per sample using the flow cytometry device MACSQuant Analyser 10 (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany) or the Attune flow cytometer (Thermo Fisher Scientific Inc., Waltham, MA). For setting the gates we used unlabelled A459 cells.

To determine the equilibrium dissociation constants (K_d) of aptamer binding to A549 cells, mean fluorescence emissions were calculated for each of seven different concentrations (50, 100, 300, 500, 800, 1,000 and 1,200 nM). The dissociation constants were calculated using one-site non-competitive binding; nonlinear curve regression was performed using GraphPad Prism 5 (GraphPad Software, Inc., San Diego, CA). Cell sorting was performed on the FACSARIA I/II equipment (Becton & Dickinson, Franklin Lakes, NJ).

Results

Preparation of aptamer pool for A549 cell line by cell-SELEX

Here we aimed to develop aptamer probes targeting NSCLC circulating tumour cells due to their potential diagnostic, prognostic and predictive capacity. For the development of NSCLC specific aptamers, we have chosen the most commonly used human lung adenocarcinoma A549 cells as target for cell-SELEX. Human blood cells were adopted as a negative control for cell SELEX to increase the selectivity of generated aptamers to A459 cells.

In the selection process, the cultured A549 cells were first incubated with a 70-base synthetic single stranded DNA library. The DNA sequences that bound to the A549 cells were then eluted and separated with the BRAZIL technique, as described in Material and Methods.

We have compared aptamer pools from the original library, the sixth cycle and the negative selection. Regarding the percentage of labelled cells for each pool, no significant difference was observed. We observed that the population of cells that were positive for the aptamer pool from negative selection had higher fluorescence, as shown by the shift to the right in the density plot and the overlap of the histograms (Figure 1A). Furthermore, we per-

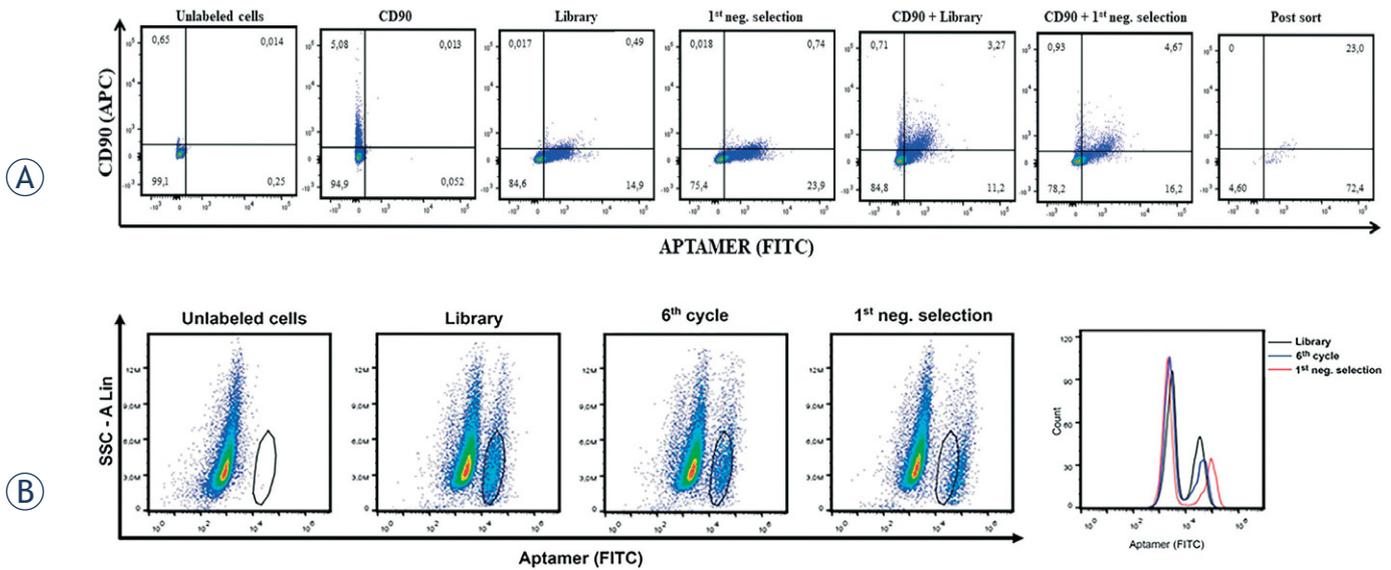


FIGURE 1. Flow cytometry of cells double labelled with CD90 and aptamers. **(A)** Flow cytometry was performed for FITC-labelled aptamer pools comparing binding to A549 cell binding of the pool of the library), the sixth cycle and the pool following negative selection against blood cells. The density plots are showing that there was no difference in the percentage of labelled cells for each of the pools. Displacement of the population in the density plot, and shift of position on the histograms for the cells labelled with the pool, which followed the negative selection against blood cells, was observed, showing a cell population with increased fluorescence intensity compared to the cells labelled with the initial library or pool of 6th cycle. **(B)** CD90 binding and cell sorting: After adjustment of cellular gates for un-labelled cells, CD90+, the library and the pool following negative selection, the double-labelled cells labelled for CD90/library and CD90/negative selection pool were compared. Here, the gate was set to collect the most intensely labelled cells with the aptamer pool following negative selection. To confirm the profile of the isolated population, the sorted cells were analysed again by flow cytometry (post sort). The density plots show the presence of a well-defined secondary population that was double labelled, suggesting that aptamers were recognizing a sub-population of CD90, which was positive for stem cells within the A549 cell population.

formed flow cytometry using antibody for CD90 detection in A549 cells and we observed a well-defined population that was positive for CD90 and aptamers (Figure 1A). This population of cells was sorted out and bound aptamers were eluted and amplified by PCR. Figure 1B reports the increase in cell labelling of the sixth selection pool following a negative selection against blood cells.

The obtained aptamer pool after the negative selection against blood cells and sorting out aptamers binding to the CD90⁺ subpopulation was used for further sequencing in order for identification aptamer candidates with conserved structural motifs.

Bioinformatics analysis

The sequencing of the seventh cycle pool by Ion Torrent PGM next generation sequencing technology resulted in 239,713 reads, from which 151,814 reads were unique. *In silico* selection of aptamers from the NGS data was accomplished using three different procedures as described in Supplementary material (Table S1).

Using CLC Genomics Workbench software, the starting pool of 239,713 reads was narrowed down

to 16 representative sequences having the highest number of sequence members, ranging from 1 080 to 142 copies per representative sequence. They were subjected to further selection based on the presence of conserved sequence motifs. Aptamers A786_1, A574_2, A278_8 were selected for *in vitro* validation, each carrying a unique motif.

In the second *in silico* aptamer selection, using online available nucleotide sequence processing programs, reads having ambiguous bases or incorrect or no sequence of the library primers were first filtered out. In the remaining high quality sequence set 80,312 reads were unique and 81.7% of the reads appeared in the set only once (no copies). Twenty sequences with the highest numbers of copies (reads), ranging from 786 to 138 copies, were ranked with regard to stability of the computed folding, using ΔG (kcal/mol) as a measure of stability, and the stability of putative loop in the three most stable computed foldings as given by the UnaFold tool. Aptamers A155_18, A452_3, A373_4, A218_12 were selected for *in vitro* validation.

In the third *in silico* selection procedure, that was executed using COMPASS software, the reads

were first clustered in 998 families with defining patterns in the loop space region. Each family was represented by a sequence (one with the highest number of repetitions) and we selected the 21 most frequent among them for further analysis. Grouping families into respective clans was depended on loop space similarity - 12 of the 21 most frequent families were positioned in 10 different family clans. Further scoring for structure stability was applied in terms of melting temperature in addition to pattern ranking and similar to the second selection procedure aptamer sequences, A155_18 and A452_3 were again selected as the best binding candidates.

The selected aptamer candidates were additionally compared with aptamer S6. The constructed dendrogram (Supplementary material, Figure S1) shows clusters of aptamer sequence candidates, where aptamer A155_18 and positive control S6 clustered together. Among the selected aptamer candidates, A155_18 and S6 are also the only two sharing two same motifs: GGTGG/CG and GCCAGT; according to the Unafold-predicted secondary structure, the motifs are placed into the comparable structural sequences between the loops (Supplementary material). The possibility of cross-contamination was excluded as the seventh SELEX pool was sequenced before the positive control S6 was purchased.

In vitro validation of selected candidates

The selected and fluorescence-labelled four candidates, A155_18, A452_3, A373_4, A218_12, together with known binder S6 and the negative control sequence (the sequence with an unknown random region with the length of 34 nucleotides, and flanked by known primers) were used in an *in vitro* binding test to determine the binding characteristics. Binding to A549 cells was demonstrated with flow cytometry analysis. All four selected aptamers (Figure 1A) showed enhanced binding to target the cell line versus the unselected control (CTRL_N34.012), according to the percentage of fluorescently labelled cells as measured by flow cytometry. Kd values of all four selected aptamers are in nanomolar range (Figure 2).

In comparison with control random aptamer, highest specific binding rates of 26% to A549 cells was obtained with A155_18. Flexible binding of aptamers at different temperatures can expand their repertoire of applications. Since the selection was performed at room temperature (25°C), we performed binding assays also at 4°C and 37°C. There

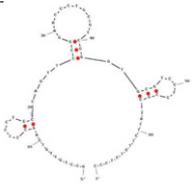
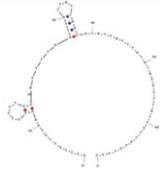
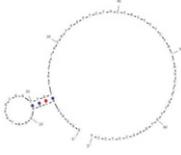
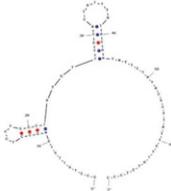
NAME	PREDICTED SECONDARY STRUCTURE	ΔG	Kd (nM)
A155_18		-5.15	160.8 ± 92.69
A452_3		-2.5	129.7 ± 81.18
A373_4		-0.36	18.65 ± 49.51
A218_12		-6.8	134 ± 123.3

FIGURE 2. Predicted secondary structures of four aptamer candidates selected for *in vitro* experiments. The secondary structures of sequences were enabling formation of complementary wrapping around their target. The unpaired bases were responsible for binding events where Watson-Crick paired bases were giving the aptamer stability. Predicted secondary structures were calculated with Unafold, the presented are the ones with lowest ΔG , i.e. the highest stability (energy rules: DNA, temperature 24°C, 14.4 mM NaCl, 3.2 mM MgCl). Equilibrium dissociation constants Kd (nM) were calculated using GraphPad Prism 5, under the non-linear fit model, one-site non-competitive binding to fluorescent population ratio at used aptamer concentrations. See the Supplementary Materials for further details.

was observed no difference in binding of aptamer A155_18 (data not shown). This is particularly important as the clinical application of aptamers will be carried out under physiological conditions.

During the cell selection aptamers are interacting specifically with the outer plasma cell membranes of targeted cells. In our cell-SELEX selection the targeted cell line A549 was enriched by the expression of the stemness marker, the membrane-bound protein CD90. This subpopulation was selected by flow sorting against CD90 in A549 cell population. The latter indicates aptamer ability of recognizing CD90-positive cells, which may be part of the tumor-initiating circulating cells.

Discussion

Cancer treatment is based on surrogate markers reflecting tumour progression and, after medical intervention, also the response to therapy. Another problem that hinders the development of new therapeutic approaches for lung cancer patients relates to the sampling of the representative (lung) tumour specimens.^{22,23,24} Circulating tumour cells (CTC) are promising as they provide an easily accessible liquid biopsy sample for real-time detection of the presence of micro-metastases. Besides their enumeration and characterization, CTC analysis offers the opportunity of mechanistic studies of malignant cancer progression⁷. It has only recently been recognized that cancer stem cells (CSC), residing in peri-arteriolar niches, undergo epithelial-mesenchymal like transition, and from there invade vascular basement membrane entering blood circulation and become CTC. In addition, a large body of evidence has accumulated on the plasticity of CSCs and the intra-tumour heterogeneity, due mainly to the presence of CSC of different subtypes and degrees of differentiation. It would thus be a nearly impossible task to target all these sub-clones by a single event. Thus, the aptamers that we developed here as biomarkers of NSCLC CTC would recognize at least one NSCLC stem cell marker, and we have succeeded in designing an aptamer recognizing a subpopulation of A549 NSCLC, expressing the CD90/Thy-1 CSC marker¹³. Indeed, following the last cycle of SELEX and sorting for CD90 positive cells, we have shown about 30 % cells are positive for this marker among the A549 cell population.

All previous aptamers developed to target NSCLC have been selected on the basis of the cell-SELEX procedure, starting with a single-stranded DNA library and proceeding with 25, 18, and 11 rounds by Zhao *et al.*, Jimenez *et al.* and Zamay *et al.*¹⁰⁻¹², respectively. All selected aptamers bind to their targets in the nanomolar range of Kd. The sequences selected by Zhao *et al.*¹⁰, the S1, S6, S11e and S15 aptamers, all had the ability to bind to A549 cells and to differentiate between normal and carcinoma lung cells. After pre-treating target cells with trypsin, the binding of selected aptamer vanished, strongly indicating that the target molecules of those aptamers were or were strongly associated with membrane proteins; however, we have not analysed them further. The target for selection used by Jimenez *et al.*¹¹ was the adenocarcinoma cell line H23, although their best binders (aptamers EJ4 and ADE1) have also shown good binding

potential for the cell line A549. However, they were not lung cancer selective, as they also recognised the colon adenocarcinoma cell line TOV21G. As possible reason for that, those aptamers recognised a more general human carcinoma cell-surface marker, of the tested cell lines were all of epithelial origin. Zamay *et al.*¹² selected DNA aptamers for lung adenocarcinoma cells derived from postoperative tissues without prior knowledge of protein biomarkers, stating that aptamers specific to the surface protein of cultured cells may not bind to tumour cells in clinical samples because of the difference in the protein expression between cultured and primary tumour cells. For cell-SELEX, these authors have used the postoperative tissue samples and have shown that the selected aptamers did not bind to healthy lung cells and the A549 lung adenocarcinoma cell line. After successful selection, they have recognized eight candidate biomarkers associated with four selected aptamers.

We have chosen the published aptamer S6, specific for A549 cells binding with low Kd value, as the positive binding control¹¹, having high selectivity for NSCLC with no binding potential against SCLC and squamous cell carcinoma cells. Its structure is similar to our selected aptamer, with 45 nucleotides in a random region, flanked by 20 nucleotide-long constant regions on both sites. As expected, aptamer S6 showed good binding potential for A549 cells also when using our protocol. The aptamer S6 was selected after 25 rounds, whereas in our SELEX procedure only seven rounds were needed. Therefore, we improved the selection process by reducing the number of selection rounds, using the bioinformatics approach. The aptamer S6 was used as a control, not only with respect to the nucleotide sequence, but also the structural motifs present, as these are crucial for binding to the selected target cell.

The *in silico* selection of aptamers was carried out after the seventh SELEX cycle, including negative selection against blood cells and yielded a satisfactory quality of sequencing reads. We were able to filter out several high-quality reads using different filters, from length to the base composition (ambiguous bases and library primers). We expected a high variability of unique sequences as there are lot of different epitopes (proteins, glycoproteins, etc.) on the A549 cells that could be potentially recognized by aptamers. The low copy number of the most abundant reads was unexpected, as for example in another cell-SELEX development, using complex bacterial spores, as SELEX targets, we obtained 10 times more copies (COBIK, un-

published). Based on the criteria used to select aptamer candidates, we suggest that physicochemical properties of computed folding (ΔG , T_m , composition of bases in the loop) are a suitable criterion to evaluate oligonucleotides obtained by only a few rounds of SELEX.

In conclusion, we have selected the aptamer A155_18, binding to a A549 lung adenocarcinoma cell line subpopulation, expressing stemness markers, such as CD90. This aptamer sequence, with two very similar motifs indicating overlapping activity with the positive control (aptamer S6), provides the proof of principle of novel approaches. Methodologically, we highly improved the reproducibility of cell-SELEX methodology when paired with bioinformatics tools. We have also shown that the use of bioinformatics reduced the number of selection cycles, thus indicating the great potential of computational biology.

Acknowledgements

We acknowledge Prof. Dr. Tanja Čufer, Dr. Ana Koren and Prof. Dr. Peter Korošec, the University Clinic of Pulmonary and Allergic Diseases Golnik, Slovenia for valuable discussions. This work was supported the Slovenian Research Agency (Program P1-0245 to TLT and the young researches fellowship to MV) and by a collaborative grant from the National Council for Scientific and Technological Development (CNPq), Brazil / Ministry of Higher Education, Science and Technology (MHEST), Slovenia, Project No. 490536/2011-5 awarded to H.U. and T.S. as well as by a grant from the São Paulo Research Foundation FAPESP (Project No. 2012/50880-4), Brazil, awarded to H.U. I.C.N. acknowledges a postdoctoral fellowship from FAPESP (Project No. 2015/18730-0). A.A.N.'s Ph.D. thesis was supported by a fellowship by FAPESP.

References

- Shigdar S, Luczo J, Wei MQ, Bell R, Danks A, Liu K, et al. Aptamer therapeutics: the 21st century's magic bullet of nanomedicine. *Open Conf Proc J* 2010; **1**: 118-24. doi: 10.2174/2210289201001001001118
- Baird GS. Where are all the aptamers? *Am J Clin Pathol* 2010; **134**: 529-31. doi: 10.1309/AJCPFU4CG2WGJJKS
- Šmuc T, Ahn IY, Ulrich H. Nucleic acid aptamers as high affinity ligands in biotechnology and biosensors. *J Pharm Biomed Anal* 2013; **81-82**: 210-7. doi: 10.1016/j.jpba.2013.03.014
- Zhou J, Rossi JJ. Cell-specific aptamer-mediated targeted drug delivery. *Oligonucleotides* 2011; **21**: 1-10. doi: 10.1089/oli.2010.0264
- Ulrich H, Wrenger C. Disease-specific biomarker discovery by aptamers. *Cytometry A* 2009; **75**: 727-33. doi: 10.1002/cyto.a.20766
- Nery AA, Wrenger C, Ulrich H. Recognition of biomarkers and cell-specific molecular signatures: aptamers as capture agents. *J Sep Sci* 2009; **32**: 1523-30. doi:10.1002/jssc.200800695
- Ni S, Yao H, Wang L, Lu J, Jiang F, Lu A, et al. Chemical modifications of nucleic acid aptamers for therapeutic purposes. *Int J Mol Sci* 2017; pii: **E1683**. doi: 10.3390/ijms18081683.
- Ulrich H, Tárnok A. Flow cytometry detection of circulating tumor cells: achievements and limitations as prognostic parameters. *Cytometry A* 2014; **85**: 201-2. doi: 10.1002/cyto.a.22441
- Krebs MG, Hou J-M, Sloane R, Lancashire L, Priest L, Nonaka D, et al. Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and -independent approaches. *J Thorac Oncol* [Internet]. 2012; **7**: 306-15.
- Zhao Z, Xu L, Shi X, Tan W, Shangguan D. Recognition of subtype non-small cell lung cancer by DNA aptamers selected from living cells. *Analyst* 2009; **134**: 1808-14. doi: 10.1039/B904476K
- Jimenez E, Sefah K, Lopez-Colon D, Van Simaey D, Chen HWI, Tockman MS, et al. Generation of lung adenocarcinoma DNA aptamers for cancer studies. *PLoS One* 2012; **7**: 1-7. doi: 10.1371/journal.pone.0046222
- Zamay GS, Kolovskaya OS, Zamay TN, Glazyrin YE, Krat AV, Zubkova O, et al. Aptamers selected to postoperative lung adenocarcinoma detect circulating tumor cells in human blood. *Mol Ther* 2015; **23**: 1486-96. doi: 10.1038/mt.2015.108
- Wang P, Gao Q, Suo Z, Munthe E, Solberg S, Ma L, et al. Identification and characterization of cells with cancer stem cell properties in human primary lung cancer cell lines. *PLoS One* 2013; **8**: e57020. doi: 10.1371/journal.pone.0057020
- Yan X, Luo H, Zhou X, Zhu B, Wang Y, Bian X. Identification of CD90 as a marker for lung cancer stem cells in A549 and H446 cell lines. *Oncol Rep* 2013; **30**: 2733-40. doi: 10.3892/or.2013.2784
- Daniels DA, Chen H, Hicke BJ, Swiderek KM, Gold L. A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment. *Proc Natl Acad Sci U S A*; **100**: 15416-21. doi: 10.1073/pnas.2136683100
- Giordano RJ, Cardó-Vila M, Lahdenranta J, Pasqualini R, Arap W. Biopanning and rapid analysis of selective interactive ligands. *Nat Med* 2001; **7**: 1249-53. doi: 10.1038/nm1101-1249
- Telez A, Rubinstein P. Rapid method for separation of blood cells. *Transfusion* 1970; **10**: 223-5. doi: 10.1111/j.1537-2995.1970.tb00733
- Nascimento IC, Nery AA, Bassaneze V, Krieger JE, Ulrich H. Applications of aptamers in flow and imaging cytometry. *Methods Mol Biol* 2016; **1380**: 127-34. doi: 10.1007/978-1-4939-3197-2_10
- Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. *Nucleic Acids Res* 2015; **43**: W39-49. doi: 10.1093/nar/gkv416
- Tamura K, Stecher G, Peterson D, Filipksi A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013; **30**: 2725-9. doi: 10.1093/molbev/mst197
- Markham NR, Zuker M. UNAFold: Software for nucleic acid folding and hybridization. *Methods Mol Biol* 2008; **453**: 3-31. doi: 10.1007/978-1-60327-429-6_1
- Kinghorn AB, Fraser LA, Tanner JA. Aptamer bioinformatics. *Int J Mol Sci* 2017; pii: **E2561**. doi: 10.3390/ijms18122516
- Berghmans T, Pasleau F, Paesmans M, Bonduelle Y, Cadranet J, Toth IC, et al. Surrogate markers predicting overall survival for lung cancer: ELCWP recommendations. *Eur Respir J* 2012; **39**: 9-28. doi: 10.1183/09031936.00190310
- Eberhard DA, Giaccone G, Johnson BE. Biomarkers of response to epidermal growth factor receptor inhibitors in non-small-cell lung cancer working group: standardization for use in the clinical trial setting. *J Clin Oncol* 2008; **26**: 983-94. doi: 10.1200/JCO.2007.12.9858