NANOBODIES AGAINST UROTHELIAL CARCINOMA

ErasmusMC
Universitair Medisch Centrum Rotterdam

Utrecht University

Erasmus Medical Center

Department of Urology

Major Research Project of the Master's Programme Science and Business Management at Utrecht University

Nanobodies against urothelial carcinoma

Bloem de Jong, BSc

Supervisors:

Joke Veldhoven-Zweistra, BSc

Prof. dr. Guido Jenster

Committee members:

Prof. dr. Raymond Schiffelers

Prof. dr. Guido Jenster

Universiteit Utrecht

Abstract

Bladder cancer (BlCa) is the 10th most common cancer worldwide, leading to more than 200,000 deaths annually. Current treatment methods of advanced disease come with many side effects and are not always effective. To improve the outcome of patients with BlCa, there is a need for more precise targeted treatments and quicker detection of primary and recurrent disease. Particularly antibodies that specifically recognize BlCa could force a breakthrough in disease imaging and treatment. Unfortunately, only very few monoclonal antibodies (mAbs) are used in the form of antibody-drug conjugates (ADCs). However, tumour accessibility is hampered due to their large size (150 kDa) and toxicity, a challenge due to limited BlCa-specificity, hence the need for smaller and more specific antibodies (Ab). Camelids produce a unique type of Ab that lack the light chain leading to a smaller Ab. When only focusing on the antigen-binding domain of the heavy chain, also referred to as the nanobody (Nb), the size is decreased to 15 kDa. In this project, we report the discovery of camelid-produced theranostic Nbs that can be used to localize and treat BlCa cells. We used M13-phage-Nb libraries, where the M13-phages display the Nb on their outer surface fused to the pIII-coat protein, to select Nbs using the phage-display technique. Firstly, we pre-selected three existing M13-phage libraries on six BlCa cell lines. Performing immunocytochemical staining of glass slides coated with multiple cancer cell lines, we showed that the cell binding affinity of these pre-selected libraries increases compared to the original libraries. Secondly, these pre-selected libraries were used to perform two consecutive rounds of biopanning against BlCa and non-BlCa cell lines. The DNA of binding M13-phage-Nbs was isolated, sequenced and used to generate a large NGS database containing the binding levels (BL) of the 300,000 most frequently sequenced Nbs that bound each individual cell line. Using this database, we selected 12 Nbs that showed high BlCa specificity and validated them using immunocytochemical staining of cells and enzyme-linked cell assays. We found three Nbs that showed high specificity towards BlCa cell lines. Also, using our NGS database, we showed the beneficial effect of two selection rounds when searching for cancer-specific Nbs, where cancer-specific Nbs increase in BL during the second selection. Although our Nbs must be validated further using patient-derived tissue material and finally within *in vivo* studies, we showed the great potential of using the phage-display technique in the search for BlCa-specific Nbs.

Keywords: Heavy chain camelid antibodies, phage-display technique, BlCa-specific Nbs, pre-selection, and biopanning.

Acknowledgements

During this project, I have learned tremendously which would have been impossible without the support from my supervisors, examiners, laboratory colleagues, fellow students, family and friends. I'm incredibly grateful and thankful that you all made my project such an enjoyable year.

First of all, I would like to thank **prof. dr. Guido Jenster** for giving me the opportunity to join the Experimental Urology Laboratory and guiding me throughout the whole project. He was available at all times for insightful feedback, and encouragement and made me grow as a scientist.

Secondly, I am profoundly grateful to my daily supervisor, **Joke Veldhoven-Zweistra.** She taught me all the skills needed to work and think independently within the laboratory. I enjoyed working together and am thankful for the freedom you gave me to shape my project.

Thirdly, I owe gratitude to **dr. Raheleh Tooyserkani** for her enthusiasm, valuable comments, knowledge and experimental assistance both in the laboratory and during the weekly nanobody meetings.

Next, I would like to thank **dr. Eric Bindels** for performing the Next-Generation Sequencing (NGS) for us, **dr. Harmen van Werken** for helping me with processing of the NGS data and **Rick Jansen** for writing the NGS data analysis scripts.

Furthermore, I would like to thank **prof. dr. Raymond Schiffelers** for being part of my thesis committee and giving feedback throughout my project.

In addition, I owe a sincere word of thanks to all other **Experimental Urology Laboratory group members** and **fellow students** for helping me around the lab, working together, giving constructive feedback, for the coffee-table conversations and the relaxing lunch breaks.

Finally, I could not have undertaken this journey without the support of my **family** and **friends**. Many thanks for their wise advice, belief and motivational pep talks.

Table of Contents

List of Figures and Tables

Abbreviations

Ab	Antibody
ADC	Antibody Drug Conjugate
AVR	Average
BL	Binding Level
BlCa	Bladder Cancer
BlCaNb	Bladder Cancer Nanobody
BlN	Bladder Normal
BSA	Bovine Serum Albumin
CCLE	Cancer Cell Line Encyclopaedia
CDR	Complementary Determining Region
DC	Dendritic Cell
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
dsDNA	double-stranded DNA
ELCA	Enzyme-linked cell assay
Fab	Fragment antigen binding domain
FBS	Fetal Bovine Serum
FFPE	Formalin-Fixed Paraffin-Embedded
FGFR3	Fibroblast Growth Factor Receptor 3
FR	Framework Region
FroCellCA	Frozen Cell Cancer Array
FroDA	Frozen cell cancer array Dab Analysis
HcAb	Heavy chain Antibody
HLA	Human Leukocyte Antigens
mAb	Monoclonal Antibody
MBS	Membrane Binding Solution
MIBC	Muscle Invasive BlCa
MOI	Multiplicity Of Infection
MWS	Membrane Wash Solution
Nb	Nanobody
NGS	Next-Generation Sequencing
NMIBC	Non-Muscle Invasive BlCa
NPs	Nanoparticles
OD	Optical Density
Ori	Origin of replication
PBMC	Peripheral Blood Monocytes
PBS	Phosphate Buffered Saline
PCa	Prostate Cancer
PCR	Polymerase Chain Reaction
PD-1	Programmed cell Death protein 1
PEG	Polyethylene Glycol
PSMA	Prostate Specific Membrane Antigen
SaaDiff	Single amino acid Difference
scFv	single chain variable Fragment
ssDNA	single-stranded DNA
STAG2	Stromal Antigen 2
SR1	Selection round 1
SR2	Selection round 2
TAA	Tumour Associated Antigen
TCR	T Cell Receptor
TERT	Telomerase Reverse Transcriptase
TMA	Tissue Micro Array
TP53	Tumour Protein 53
TSA	Tumour-Specific Antigen
TUR	Transurethral resection
RPMI	Roswell Park Memorial Institute
RT	Reverse Transcriptase
VHH	Variable Heavy Chain
WT	Wild type

1 Theory

1.1 Anatomy of the bladder

The human body consists of multiple biological systems that work together to maintain the functions that are needed for movement, reproduction, metabolism, response, organization, growth, etc. All biological systems function within a precisely-tuned balance and can become disrupted in the case of diseases, such as cancer. An example of a biological system is the urinary system, which is needed to produce, store and excrete urine that contains harmful molecules that must be removed [1]. Organs belonging to the urinary system are the kidneys, renal pelvis, ureters, bladder and urethra [2]. The bladder stores urine and secretes urine through the urethra when the bladder is about 25% full [1]. On a cellular level, the bladder consists of four different layers of cells: urothelium, lamina propria, muscularis propria and serosa (Figure 1) [3]. Going from the interior of the bladder to the exterior, we first encounter the urothelium, a three to seven cells thick layer that separates urine from the muscle tissue. This layer requires constant maintenance and repair, and consists of three different cell types: the basal cells which play a role in regeneration of the urothelium, the intermediate cells which regenerate tissue (for example in case of infection) and the superficial layer which forms a strong barrier for water and solutes [3]. Secondly, the lamina propria layer is considered the communication centre of the bladder and plays an important role in signal transduction, both locally between the different cell layers via the extracellular matrix and distantly to the central nervous system via local sensory nerves [4]. Thirdly, the muscularis is, as the name suggests, a layer consisting of muscle cells, needed for contraction and controlled expansion of the bladder. Lastly, the serosa layer covers the outer part of the bladder with a thin layer of connective tissue.

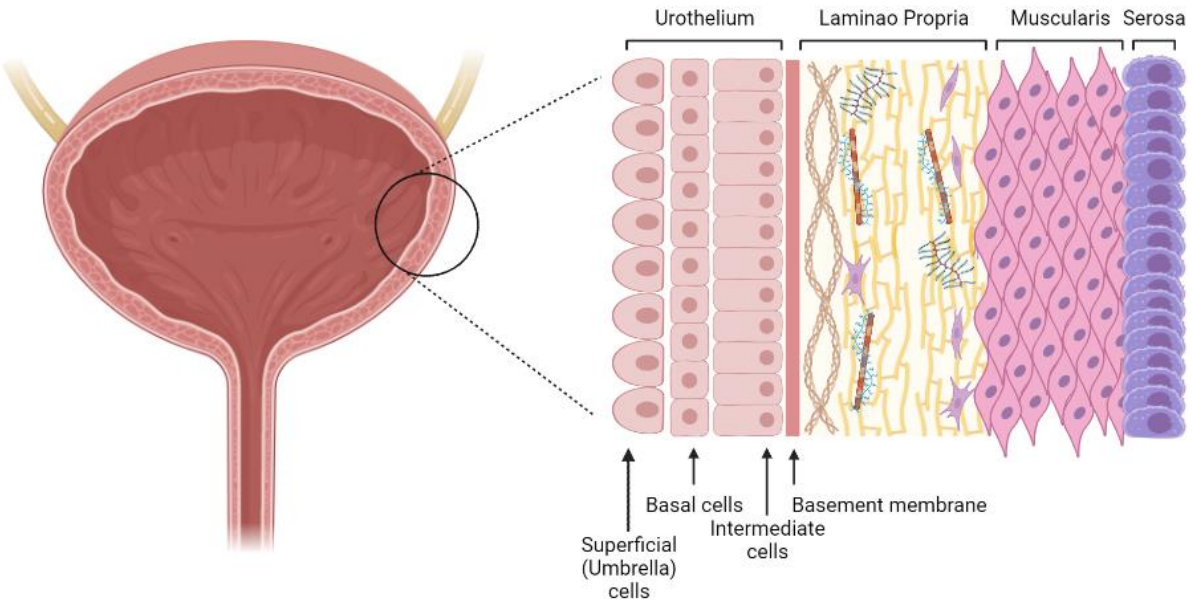


Figure 1 | Anatomy of the bladder showing the four different cell types which construct the bladder wall. Also, the two ureters and the urethra are shown. *Adapted from Khandelwal et al. (2009)* [5] and *created with BioRender.com.*

1.2 Bladder cancer

Bladder cancer (BlCa) starts at the cellular level and can be seen as a multistep process. It involves mutations and the selection of cells that can evade growth suppressors, sustain proliferative signalling, resist cell death, induce angiogenesis, enable replicative immortality and in more progressed stages of cancer even activate invasion and metastasis [6]. These cancer cells gradually become the masters of their progression through progressive series of alterations and in the end develop to a stage of malignancy [7-8]. BlCa accounts for 3% of all cancer diagnoses and is the 10th most common cancer worldwide [9]. In 2020, 533,278 new cases of BlCa were reported of which 212,536 patients did not survive, leading to a mortality rate of approximately 40% [9]. Additionally, it was reported that approximately 80% of BlCa cases occur in males.

Mostly, BlCa starts in the urothelium and when left untreated might invade deeper into the bladder tissue (Figure 2). Tumours that grow into the muscularis are considered muscle-invasive bladder cancers (MIBCs) and are known to spread quickly to other parts of the human body [10], often to nearby lymph nodes, bones, lungs and the liver [11]. For more deeply invaded tumour tissue, the survival rate of patients decreases rapidly. For grade 1 tumours (Tis, Ta and T1), the five years survival rate is 80%, decreasing to 45% for grade 2 tumours (T2a, T2b), to 40 % for grade 3 tumours (T3) and even to 10% for grade 4 tumours (T4) [12].

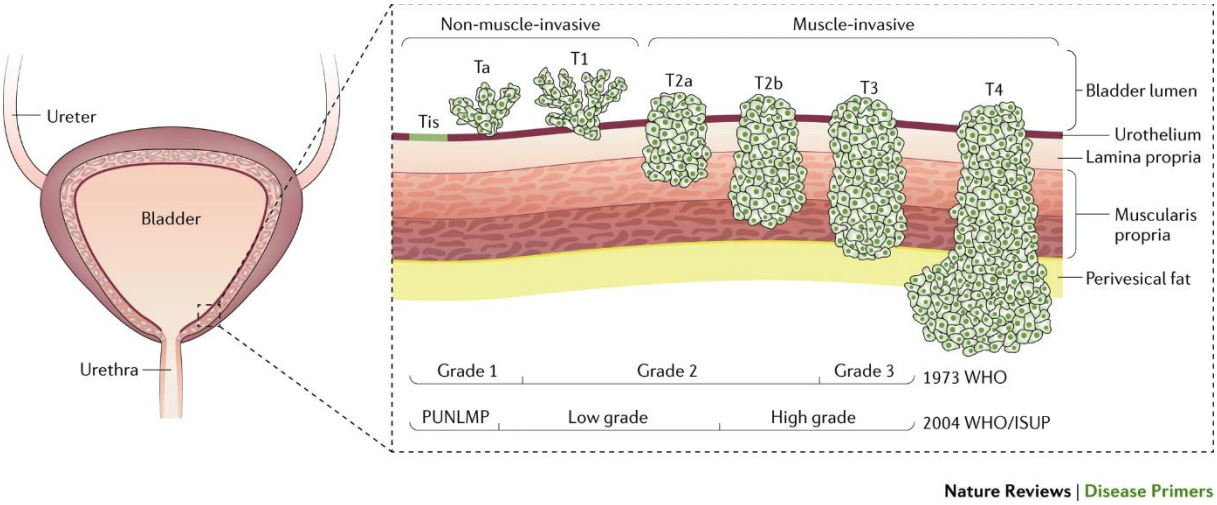


Figure 2 | Staging of BlCa. Ta and T1 are considered non-muscle-invasive cancer, while T2a-T4 are considered muscle invasive. Stage T2a has superficially invaded the muscle layer, while T2b invaded more deeply. T3 has invaded the perivesical fat and T4 is metastatic cancer that has spread to other organs. *Adapted from Sanli et al. (2017)* [10].

Looking at the pathological mechanisms that underly BlCa, two distinct pathways can be indicated, one giving rise to non-muscle invasive BlCas (NMIBCs) and one to MIBCs. For example, NMIBCs are characterized by loss of heterozygosity of chromosome 9, activating mutations of genes that encode for fibroblast growth factor receptor 3 (FGFR3), telomerase reverse transcriptase (TERT) and mutations in genes such as PIK3CA and stromal antigen 2 (STAG2). These mutations all lead to superficial tumours, due to excessive cell proliferation, division and growth. The other pathway that leads to MIBCs is

characterized by having tumour protein 53 (TP53) mutations and shows a loss of heterozygosity of chromosome 9 [10, 13-15].

1.3 Possible therapies

Currently, when BlCa is diagnosed, there are multiple treatments possible. The first one is surgery, where the tumour is removed using transurethral resection (TUR). Secondly, when the cancer is more invasive, one can perform a radical cystectomy to remove the bladder completely. In some cases, cancer has also spread to nearby organs, such as the uterus in women or the prostate in men, which is then almost always removed, independent of the cancer stage. When BlCa is located in one part of the bladder and is minimally invasive, surgery can also be used to remove only part of the bladder. Other therapies that do not involve surgery are radiation, chemo-, and immunotherapy. With radiation, high-energy x-rays are used to kill cancer cells and prevent them from growing, whereas chemotherapy drugs are used to kill cancer cells. An example of BlCa-directed chemotherapy is enfortumab vedotin-ejfv [16], which is a Nectin4-directed antibody-drug conjugate (ADC). With immunotherapy, the patient's immune system is stimulated to recognize cancer cells and kill them [1]. An upcoming new immunotherapy is dendritic cell (DC) vaccination, in which DCs uptake the antigens of cancer cells and present them to the patient's T cells which leads to activation of the patient's immune system. Another example of immunotherapy is pembrolizumab, which is an antibody (Ab) targeting programmed cell death protein 1 (PD-1) [17].

Unfortunately, current treatment methods are not very effective and typically only delay survival rate on average by months to a few years. There is a need for new therapies that are more precisely targeted and lead to fewer side effects for the patients. Targeted nanomedicine has been shown to improve efficacy for subgroups of patients and can be less damaging for healthy cells surrounding cancer cells. They allow targeting of specific membrane proteins present on the outer surface of cancer cells (e.g. Nectin4, PD-1) and can interfere with the signal transduction of the cancer cells.

1.4 Nanobodies

Cancer cells can obtain specific types of membrane proteins, namely tumour antigens, which serve as a biomarker for the identification of cancer cells. These antigens may be used to develop new therapeutic agents to trigger the immune response against these cancer cells or even kill cancer cells directly by the delivery of anti-cancer drugs [18].

A specific class of proteins produced by the immune system, that recognize antigens on foreign substances, are called antibodies (Abs). Monoclonal antibodies (mAbs) (Figure 3A) are selected and produced in the laboratory and can be used to inhibit tumour cell proliferation and interfere with cancer cells by targeting antigenic membrane proteins [19]. Regular Abs consist of two heavy and two light chains, which form two identical fragment antigen binding domains (Fabs). Antigen specificity is determined through complementary determining regions (CDRs), where CDR3 is the most important in

generating diversity and specificity among mAbs. Unfortunately, due to their large size (150 kDa), tumour accessibility is hampered. The structural properties of the antibody recognition module lead to difficulties when producing bispecific agents as the CDR is determined by two different protein chains. Another drawback of these mAbs is their relatively long half-life, which leads to large background signals when used for medical imaging [20-21]. Most of the disadvantages of mAbs are due to their large size and hence the need for new, smaller, single-chain antibodies with high stability *in vivo*.

Heavy-chain only antibodies (HcAbs) (Figure 3A), lacking the light chain, produced by camelids show great prospects as anti-tumour therapeutic agents [22]. These antibodies are smaller (~ 74 kDa) than conventional antibodies and the antigen binding domain or so-called variable heavy chain (VHH) domain is only 15 kDa in size and is often referred to as the 'nanobody' (Nb) [23]. This VHH domain contains three CDRs, where the CDR3 domain is responsible for 60-80% of the binding interaction with antigens [24-25]. The CDR3 domain on HcAbs has an increased length compared to mAbs, leading to potentially higher binding affinities and the ability to bind cavity-like domains on cancer cell membrane proteins [26]. The lack of the light chain also leads to lower immunogenicity and increased stability [27-28]. Another important feature of these nanobodies, due to their small size, is their ability to cross the blood-brain barrier [29].

In many cases, tumour-associated antigens are also expressed by healthy cells, causing the Nb to bind to both healthy and cancer cells. This may be circumvented with the use of bispecific Nbs (Figure 3B) that bind two distinct tumour antigens. However, the downside of these dimeric Nbs is again the increased size, leading to reduced tissue penetration. Nbs can also be used as targeting moieties of effector domains, such as cytokines that trigger the immune response or as antibody-drug conjugates (ADC) that deliver immunotoxins (chemotherapy agents) to cancer cells (Figure 3B). Another implementation of Nbs is to use them as a targeted delivery system of nanoparticles (NPs) (< 200 nm) with encapsulated drugs (Figure 3B). However, with such NPs, the stability of the Nbs might be attenuated and the size might again lead to disadvantages. By using human leukocyte antigens (HLA) - specific Nbs on the outer surface of the NPs, we could prevent immune recognition by the host's leukocytes and therefore prevent rapid elimination from circulation [30]. Finally, Nbs can be used as imaging probes (Figure 3B), for example by attaching fluorescent dyes or radionuclides to the Nbs.

Taken together, Nbs offer attractive properties to improve cancer therapy either by molecular conjugation or coupling to larger nanocarriers. These Nbs have favourable properties compared to conventional mAbs and during this project, we have exploited the potential of antigen-specific Nbs where we aim to target BlCa antigens.

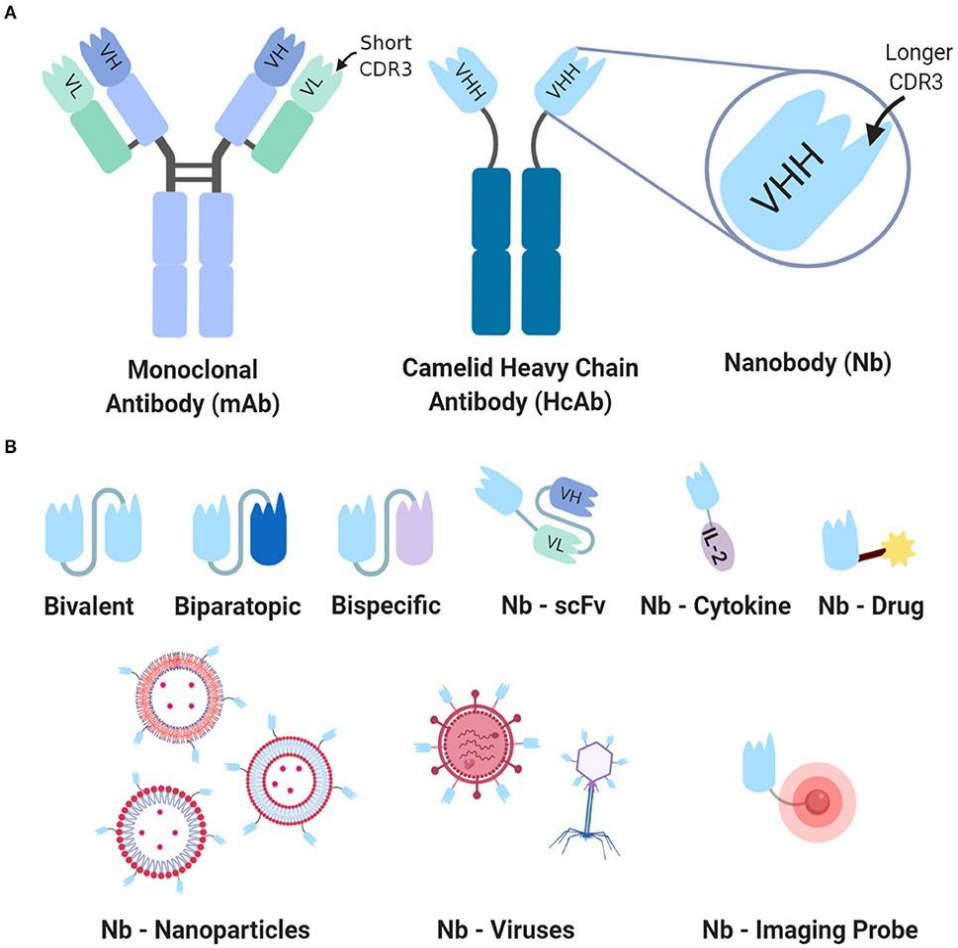


Figure 3 | Monoclonal and heavy chain antibodies. A (left) A schematic image of a monoclonal antibody (mAb) showing both the light and heavy chain including the short CDR3 domain. (right) A schematic image of a camelid heavy chain antibody (HcAb) showing the heavy chain constructing this antibody which conducts a longer CDR3 domain. **B** Multiple applications of the VHH chain (Nb) from the HcAb of camelids. These Nbs can be modified to become bivalent, biparatopic, bispecific, or can be combined with other components to become Nb-scFv, Nb-cytokine or Nb-Drug fusions. Nbs can also be used as nanoparticles, Nb-virus fusions or as imaging probes. *Adapted from Yang et al. (2020)* [22].

1.5 Phage-display technique to select for nanobodies

Nanobodies of interest can be selected using a technique called phage-display where nanobodies are fused to the coat proteins of the phage and are 'displayed' to the external milieu. In this project, we used the M13 phage which consists of five constructive proteins (Figure 4), has an ssDNA genome packed inside and use *Escherichia coli* (*E. coli*) as their host bacteria. With the phage-display technique, the gene encoding for the Nb of interest is included in a dsDNA-phagemid vector and fused with the pIII coat protein gene, leading to a fused pIII-Nb protein 'displayed' on the outer surface of the M13 phage. The pIII protein also plays a major role in the M13 life cycle. Here, pIII interacts with the F pilus, a membrane protein displayed on the outer surface of *E. coli*, to initiate cell entry. Upon entry, the ssDNA from the phage is converted into dsDNA using the host DNA replication machinery [31]. Infection with M13 phages does not kill the host cell, rather the M13 phage uses *E. coli* to replicate, assemble, multiply and formation of new phages [31-36].

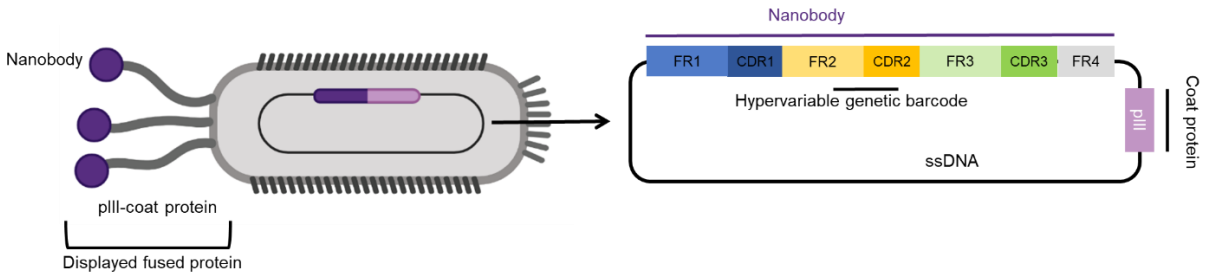


Figure 4 | Schematic representation of the M13 bacteriophage. Showing the ssDNA containing the Nb sequence (CDR1-3, FR1-3) consisting of all CDR regions interspersed by framework regions (FR) and the pIII gene. *Created with BioRender.com.*

1.6 Nanobody selection

We used large M13-phage libraries constructed from camelid blood samples for the selection of BlCa-specific Nbs. Three libraries (L1P4, LUPCa1, LUPCa2) were generated by the immunization of three llamas. Libraries LUPCa1 and LUPCa2 were selected by immunization with a mix of isolated single cells from frozen prostate cancer (PCa) and BlCa tissue, whereas for L1P4 only PCa tissue was used. After immunization, the RNA was isolated from isolated peripheral blood monocytes (PBMCs) of the llamas. Using reverse transcriptase (RT), cDNA was made from this RNA and the VHH chain was isolated using polymerase chain reaction (PCR) with a specific set of primers. With the help of restriction enzymes SfiI and NotI, the Nb-fragments are cloned into the M13 phage genome at the 5' end of the pIII gene, leading to a fused pIII-Nb protein when expressed. L1P4 was an already existing library, while libraries LUPCa1 and LUPCa2 were constructed by the Experimental Urology Laboratory located at Erasmus Medical Center. Using these libraries they started the search for cancer-specific nanobodies.

The ultimate goal would be to find a specific Nb that binds only to cancer and not healthy cells. To select these Nbs, biopanning is used, which is an affinity selection technique used to select Nbs that bind cancer cells. With biopanning, phages expressing pIII-Nb fused proteins are incubated with specific cancer cell types where phages that can bind to membrane proteins present on the cancer cells remain. The sample is washed multiple times to remove unbound phages and the bound phages need to be re-grown using *E. coli* as a host. Selected phages can infect new bacteria from which the phagemids can be isolated to find the selected Nb sequences. To perform a consecutive round of biopanning, *E. coli* are re-infected with the recovered phages containing the phagemid DNA with only the pIII-Nb gene, but not the genes needed for phage-assembly, chromosome replication and amplification [37]. Therefore, we need a helper phage, for example, M13K07Δ3-pIII, that will complete the assembly of new phages containing a fused pIII-Nb coat protein. The *E. coli* produce new phages that are selected for specific cancer cell types and this new phage library can be used for another biopanning round, further selecting for binding of Nbs step by step. After each biopanning round, we isolated the DNA of all samples against which panning was performed and used this DNA for next-generation sequencing (Figure 5).

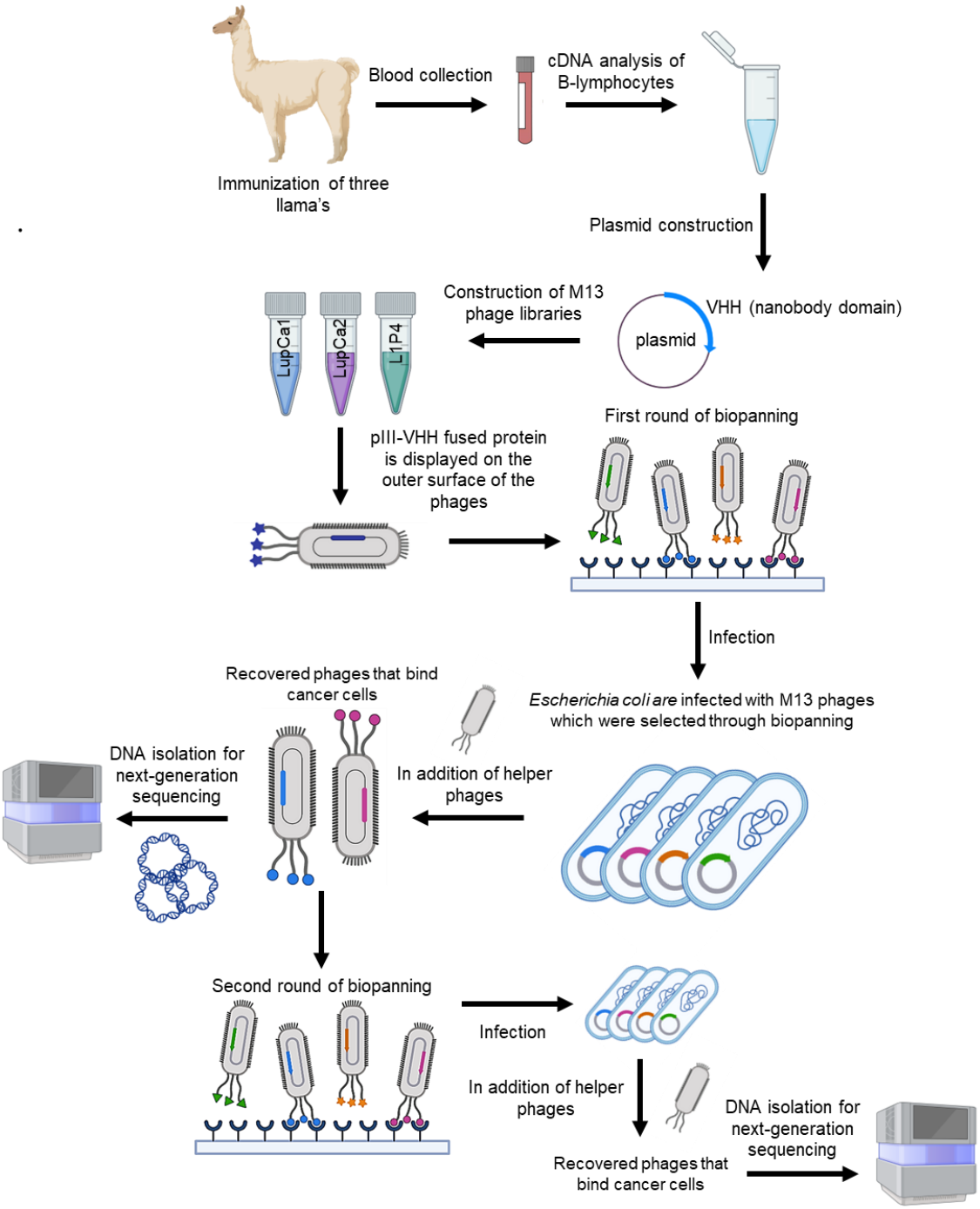


Figure 5 | Schematic representation of the biopanning and production of new phages. M13 phage libraries are incubated with cancer cells, leading to the selection of M13 phages that can bind cancer cell proteins. These phages are used to re-infect *E. coli*. With the helper phages, new phages are produced and used for the second round of biopanning. *Created with BioRender.com.*

The Experimental Urology Laboratory has found multiple Nbs using this technique. For example, JVZ-007, also referred to as A7, which is a prostate-specific membrane antigen (PSMA) binding Nb [38]. Other Nbs found are C5 and H6, which bind to CD63 and CD9 proteins. Both proteins are associated with cell differentiation, cell fusion and cell migration [39]. Here, the tetraspanin CD9 is a protein that is expressed on the surfaces of leukocytes (B cells, T cells, granulocytes, etc.), at high levels on endothelial cell surfaces and is found on almost all cells and many of the extracellular vesicles [40].

CD9 has been shown to have characteristics similar to tumour suppressor proteins and low levels of CD9 are associated with lower survival rates [41]. Also, CD9 is shown to be a biomarker for metastatic clear cell renal cell carcinoma and therefore exhibits both protumourigenic and tumour suppressor properties [42-43]. Tetraspanin CD63 is a protein mostly expressed in endosomes and exosomes. It plays a role in platelet activation [44] and is associated with Hermansky-Pudlak syndrome [45]. CD63 was reported to be upregulated in the early stages of breast cancer, astrocytoma and melanoma [39], while being downregulated with higher invasiveness of tumours [46-48]. These three Nbs prove that our adapted phage-display technique can be used to find specific Nbs against membrane proteins and form the basis for the search for BlCa Nbs.

2 Research objective and workflow

The experimental urology research group uses the M13 phage-display technique to rapidly select Nbs that specifically recognize cancer cells while not binding to normal cells. The binding potential of billions of nanobodies against cancer and normal cells is collected and organized in their proprietary database. Using this database, they were able to select PCa-specific Nbs that turned out to bind the well-known disease target prostate-specific membrane antigen (PSMA) [38]. With this proof of concept, our overall goal is to optimize the panning procedure and identify novel BlCa-specific Nbs. We will utilize the M13-phage libraries L1P4, LUPCa1 and LUPCa2 for this goal.

This project can be described by the following steps (Figure 6):

1. **Library pre-selection**: Panning of the libraries L1P4, LUPCa1 and LUPCa2 against a collection of BlCa cell lines to create a compiled targeted sub-library against BlCa cells.
2. Use this pre-selected Nb-phage library for **two rounds of panning** against different cancer cell lines (BlCa, PCa and other Ca's).
3. Perform **Next-Generation Sequencing** of the bound phages to each of the tissues and compile the Nb sequences in a database.
4. **Select nanobodies** with a BlCa-specific binding profile.
5. **DNA synthesize Nbs as cDNA** into a phagemid by the company Twist Bioscience.
6. Followed by the transformation of the phagemids in bacteria, **generate batches of Nb-phages**.
7. **Test Nb-phage binding** to different cell lines for validating their specificity.

In this project, we specifically **aim to gain more understanding of the effect of multiple selection rounds and the pre-selection of phage libraries** when searching for BlCa-specific Nbs. The specificity of selected Nbs and the functionality of our Nb NGS database will be validated using adapted assays for Nb binding to cell lines and tissues.

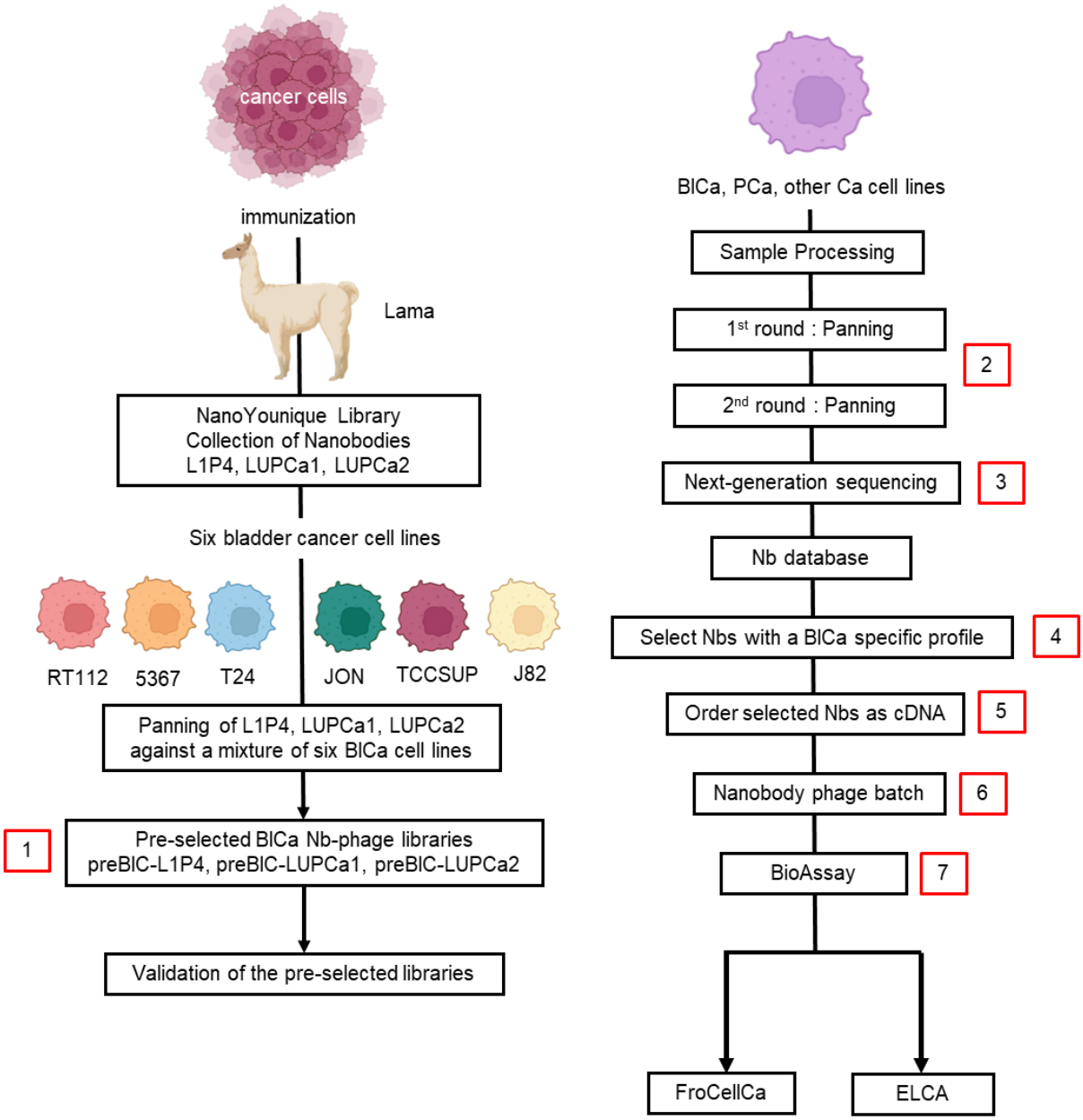


Figure 6 | Flow scheme showing the project outline. 1 = Library pre-selection, 2 = two rounds of panning, 3 = next-generation sequencing, 4 = select Nbs with a BlCa specific profile, 5 = order selected Nbs as cDNA, 6 = nanobody phage batch generation, 7 = bioassays to test Nb binding. *Created with BioRender.com.*

3 Materials and Methods

3.1 Cell culturing

All cell lines used were cultured in Dulbecco's Modified Eagle Medium (DMEM) or Roswell Park Memorial Institute (RPMI) 1640 medium with 10% fetal bovine serum + P/S (FBS). We used T25, T75 and T175 flasks for adherent cell lines and split cells at a 1:4 to 1:8 ratio. When splitting the culture, first the medium was removed and cells were washed with phosphate-buffered saline (PBS): 4 mL for T25, 10 mL for T75 and 16 mL for T175 flasks. PBS was removed and trypsin (1, 2 or 4 mL) was added and incubated at 37°C (to detach the cells from the flask) until all cells were fully floating through the flask. The cells were resuspended in medium and centrifuged at 1200 rpm for 5 min (Eppendorf centrifuge 5804 R). Finally, the cells were resuspended in an appropriate volume medium (4, 12 or 22 mL) and plated into a new flask.

3.2 Pre-selection of L1P4, LUPCa1 and LUPCa2 on BlCa cell lines

We pre-selected three M13-phage libraries: L1P4, LUPCa1 and LUPCa2, on six different BlCa cell lines (RT112, T24, TCCSUP, 5637, JON and J82). During this protocol, the three libraries were kept separated and after pre-selection the libraries were named preBlC-L1P4, preBlC-LUPCa1 and preBlC-LUPCa2. For the following protocol, we edited and used the CESAME Manual Version 2.0 *(1996)* [49].

Step 1: BlCa cell mixture preparation

We used 15 x 10^6 cells for each cell line, leading to a total of 90 x 10^6 used for pre-selection. The cell lines were stored at - 80ºC with dimethyl sulfoxide (DMSO) to prevent the formation of ice crystals when becoming frozen. For cell preparation, we performed the following steps:

1. The cells were melted at room temperature.
2. All six BlCa cell lines were mixed.
3. We added RPMI-FBS 10% medium to dilute the cells 10-fold.
4. The cells were centrifuged at 1,200 rpm for 5 min (Eppendorf centrifuge 5804 R).
5. The supernatant was removed.
6. The pellet was resuspended in 15 mL 1% BSA (bovine serum albumin)-$_{1x}$PBS(137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 and 2 mM KH_2PO_4) and centrifuged again for 5 min at 1,200 rpm (Eppendorf centrifuge 5804 R).
7. The supernatant was removed and finally, the cell pellet was resuspended in 3 mL 1% BSA-$_{1x}$PBS.

Note: Cells were detached from the flasks before storage with 3.4 mM EDTA in PBS, not using trypsin. Since trypsin may cleave membrane proteins.

Step 2: Phage library preparation

Library L1P4 has 3.0 x 10^9 different phages and therefore we added 3.0 x 10^{12} phages to the BlCa cell mixture. Phages were diluted in 1% BSA-$_{1x}$PBS, mixed gently and incubated for 15 min (Table 1) to block unspecific binding.

M13 libraries	Library size (independent phage clones)	Total number of phages added to the BlCa cells
L1P4	3.0 x 10^9	3.0 x 10^{12}
LUPCa1	1.2 x 10^9	1.2 x 10^{12}
LUPCa2	8.0 x 10^8	8.0 x 10^{11}

Table 1 | M13 phage library size and calculation of the number of phages needed to reach 1000-fold excess of phages.

Step 3: Mixing of phages with BlCa cells

We incubated the BlCa-cell mixtures with either L1P4, LUPCa1 or LUPCa2 at room temperature for ~ 60 min while being gently mixed on a Heidolph duomax. After incubation, the mixtures were centrifuged at 1600 rpm for 5 min (Eppendorf centrifuge 5804 R). The supernatant was stored and contains unbound phages. The pellet contains BlCa cells with bound phages which will be washed in *step 5* of this protocol.

Step 4: TG1 culture

Simultaneously with the incubation of phages and BlCa cells, the TG1 culture was grown in 2xTY (5 grams of NaCl, 10 grams of yeast extract and 16 grams/L Casein) where we added 0.5 mL (1:100) of an overnight culture of *E. coli* TG1 to the 2xTY while being shaken at 37ºC for ~105 min at 270 rpm (New Brunswick Scientific C25 Incubated Floor Model Shaker). The *E.coli* grow exponentially and must be used immediately as they express their F-pilus only in the exponential growth phase (optical density at 600 nm (OD_{600}) of 0.5-0.7). This F-pilus is needed for infection with phages.

Step 5: Removal of unbound phages

We washed the pellet from *step 3* and removed unbound phages from the cell/phage mixture. We added 30 mL 0.5% BSA-$_{1X}$PBS to resuspend the pellet and centrifuged the mixtures at 1,600 rpm for 5 min (Beckman centrifuge J-6B). Next, the supernatant was removed and the cells were resuspended in 30 mL 0.5% BSA-$_{1X}$PBS (mix gently). These washing steps were repeated 5-10 times. During washing, the cells were transferred to fresh tubes twice. Finally, the supernatant was removed, leaving us with a pellet of cells at the bottom of the tube.

Step 6: Elution

The pile of cells was resuspended in 2.6 mL H_2O (0.6 mL per 21 x 10^6 cells), followed by dropwise addition of 2.6 mL 200 mM trimethylamine (TEA) while vortexing and incubated for 5 min. Next, 2.6 mL of 1M Tris-HCl (pH 7.4) was added dropwise while vortexing, leading to a final volume of 7.8 mL. The mixtures were centrifuged at 14,000 rpm for 5 min (Eppendorf centrifuge 5804 R). The supernatant (containing preBlC-selected phages = phage eluate) was transferred to a fresh tube.

Step 7: Re-infection

For re-infection, the TG1 culture was mixed with the phage eluate from step 5. We added TG1 culture to the phage eluate (10:1 ratio) to dilute the TEA ten-fold and incubated without shaking for 45 min at 37ºC. During this phase, the bacteria get infected with phage DNA. After infection, the bacteria were centrifuged for 5 min at 4,200 rpm (Beckman centrifuge J-6B). Next, the bacteria were resuspended in 2xTY with ampicillin (100 µg/mL) and glucose (2%). Ampicillin kills bacteria that do not contain our phagemids as these contain a gene domain that causes bacteria to become resistant to ampicillin. Finally, we incubated overnight in 800 mL TAG at 30ºC while being shaken at 270 rpm (New Brunswick Scientific C25 Incubated Floor Model Shaker). Important to always use conical flasks at least 5x bigger than the volume used and to keep the libraries separated from one another.

Step 8: Titration

To estimate the sizes of the pre-selected BlCa-libraries, we saved 5 µL of the phage eluate of each library. We made serial dilutions of the phage supernatant as follows:

- 10^{-2} dilution = 5 µL of phages with 495 µL PBS
- 10^{-4} dilution = 5 µL of the 10^{-2} dilution with 495 µL PBS
- 10^{-5} dilution = 50 µL of the 10^{-4} dilution with 450 µL PBS
- 10^{-6} dilution = 50 µL of the 10^{-5} dilution with 450 µL PBS

To the 10^{-5} and 10^{-6} dilutions, we added 500 µL of our exponentially growing TG1 culture and incubated for 30 min at 37 °C. We plated 100 µL of the 10^{-5} and 10^{-6} dilutions on agar plates containing 2% glucose and 100 µg/mL ampicillin. Finally, the plates were incubated overnight at 37°C. The next day, the colonies were counted and used to estimate the retrieved phages. The plates were stored at 4°C.

> *For example,* we count 50 colonies on the 10^{-6} plate. The library size then has a size of 50 x 10^6 x 2 (1:1 diluted in TG1) x 10 (as 100 µL was plated), leading to 1 x 10^9 different phages.

Step 9: Next steps

After the overnight incubation, the bacteria cultures were distributed over 50 mL tubes. These bacteria were used for:

1. Storage by making glycerol stocks (see step 10)
2. Phage production (see 'Materials and Methods' section 3.3)
3. DNA isolation (see 'Materials and Methods' section 3.5)

Step 10: Glycerol stocks

First, the bacteria cultures were centrifuged at 4,200 rpm for 15 min (Beckman centrifuge J-6B). The supernatant was removed and the pellet was resuspended in 600 µL 2xTY. Secondly, 300 µL sterile glycerol (at 80% concentration) was added to the bacteria leading to a final 25% glycerol concentration. Glycerol stocks were stored at -80ºC.

3.3 Pre-selected phage library production

For the following protocol, we edited and used the CESAME Manual Version 2.0 *(1996)* [49]. The pre-selected libraries were named **preBlC-L1P4**, **preBlC-LUPCa1** and **preBlC-LUPCa2**.

Step 1: Calculations of the starting culture

Using the library sizes calculated after titration (see 'Materials and Methods' 3.2 step 8), we estimated that the libraries have been enriched 20-fold. Therefore, we have used the following library sizes as starting conditions: preBlC-L1P4 (1.50×10^8 phages), preBlC-LUPCa1 (6.00×10^7 phages) and preBlC-LUPCa2 (3.00×10^7 phages). The inoculum size of the bacteria should be added to the libraries with a multiplicity of infection (MOI) of 100 to the library size (Table 2) and should not exceed 0.08 OD_{600nm}. An OD of 100 is similar to a concentration of 8.0×10^{10} bacteria/mL. See *for example* calculation for preBlC-L1P4 and apply this to all library calculations.

> *For example*, when measuring the stock of bacteria infected with the preBlC-L1P4 library, we measured an OD of 1.2, meaning we had 7.68×10^9 bacteria/mL. For preBlC-L1P4, we needed to start 1.50×10^{10} bacteria. Therefore, we needed to take 1.95 mL of the preBlC-L1P4 stock and diluted it in 234 mL medium (2x TY, 100 µg/mL ampicillin and 2% glucose).

	Initial library size	Starting library size preBlC-selection	Number of bacteria
L1P4	3.00×10^9	1.50×10^8	1.50×10^{10}
LUPCa1	1.20×10^9	6.00×10^7	6.00×10^9
LUPCa2	6.00×10^8	3.00×10^7	3.00×10^9

Table 2 | Number of bacteria needed to produce for whole coverage of library diversity assuming that the L1P4, LUPCa1 and LUPCa2 libraries have enriched 20x during pre-selection.

Step 2: Grow up new bacteria expressing the F pilus

The preBlC-infected bacteria from step 1 were incubated for approximately 2 to 3 hours at 37ºC while being mixed at 270 rpm. The OD_{600nm} was measured every 30 min until an OD_{600nm} of 0.6-0.8 was reached.

Step 3: Addition of the helper phage

When an OD_{600nm} of 0.6-0.8 was reached, we added the helper phages to the bacteria. We added the helper phage with a MOI of 5 to the bacteria. We used the M13K07Δ3-PIII helper phage (Progen) which is a pIII defective helper phage. The mixture was incubated for 30 min at 37ºC without shaking, followed by 30 min at 37ºC at 120 rpm (New Brunswick Scientific C25 Incubated Floor Model Shaker) (only in the case of large libraries).

> *For example*, if the bacteria have an OD_{600nm} of 0.8, we have 6.40 x 10^8 bacteria/mL and aim for 5 x 6.40 x 10^8 = 32 x 10^8 helper phages / mL. In the case of a helper phage stock concentration of 2.3 x 10^{12} phages/mL, we add 719 µL helper phages.

Step 4: Grow phages overnight

The bacteria were centrifuged for 15 min at 4,000 rpm (Beckman centrifuge J-6B). The supernatant was removed and the bacteria pellets were resuspended in 2x TY with 25 µg/mL kanamycin and 100 µg/mL ampicillin. In this case, next to ampicillin which kills uninfected bacteria, also kanamycin was added to the medium which will kill bacteria that do not contain the helper phage (vector). The helper phage vector contains a gene that leads to kanamycin resistance. The resuspended bacteria were grown overnight at 30ºC at 270 rpm (New Brunswick Scientific C25 Incubated Floor Model Shaker). Again, it is important to always use conical flasks at least 5 times bigger than the volume used.

Step 5: Precipitation of the phages

The next day, the bacteria were centrifuged for 15 min at 4,000 rpm (Beckman centrifuge J-6B) and resuspended the bacteria in PBS. We added $\frac{1}{5}$ of the total end volume polyethylene glycol $(PEG)_{6000}$ / 2.5 M natrium chloride (NaCl) to the supernatant, mixed well and left the sample on ice for 90 min. The PEG_{6000} causes the phages to separate from the medium. Next, the samples were centrifuged for 30 min at 4,000 rpm and the supernatant was removed well by decanting and then dripping on tissue with aluminium foil underneath. Again, each pellet was resuspended in PBS, we centrifuged the mixture for 30 min at 4,000 rpm, we added $\frac{1}{5}$ of total end volume PEG_{6000} / 2.5 M NaCl to the supernatant, mixed well and left the sample on ice for 45 min. We centrifuged the sample for 30 min at 4,000 rpm and discarded the supernatant. Again, we resuspended the phages collected in the pellet in 10 mL PBS and divided each library into five Eppendorf tubes of 2 mL. We centrifuged all the tubes for 15 min at 10,000 rpm (Eppendorf centrifuge 5417R). Finally, we pressed the supernatant through 0.45 µm filters and

stored the phages (clear solution) at 4 °C. As mentioned before, the pre-selected libraries were named **preBlC-L1P4**, **preBlC-LUPCa1** and **preBlC-LUPCa2**.

Step 6: Phage concentration calculations using Nanodrop

The stock concentration of the phage libraries can be measured quantitatively using Nanodrop (Thermofisher scientific Nanodrop ND-1000). We made serial dilutions of the preBlC-libraries (5x, 10x, 20x, 40x and 80x dilutions) and measured the OD of each dilution at 268 λ. For the standard curve, we made a serial dilution of the helper phage which has a known concentration of 1.1 x 10^{12} phages/mL. Using this standard curve, we calculated the concentration of each preBlC-library, leading to concentrations (phages/mL) of preBlC-L1P4 = 1.4 x 10^{12}, preBlC-LUPCa1 = 6.9 x 10^{11}, preBlC-LUPCa2 = 2.1 x 10^{12}.

3.4 Panning of pre-selected libraries on Ca cell lines

After the precipitation of the pre-selected phages, we used the preBlC-L1P4, preBlC-LUPCa1 and preBlC-LUPCa2 to perform two consecutive panning rounds against Ca cell lines. We have used 10^7 cells for each panning. After each panning, we isolated the DNA (see 'Materials and Methods' section 3.5) which will be used for polymerase chain reaction (PCR) and next-generation sequencing (NGS). For the following protocol, we edited and used the CESAME Manual Version 2.0 *(1996)* [49].

Note: For the following protocol, we only mentioned adjustments made to the protocol described in sections 3.2 and 3.3.

3.4.1 Selection of phages – round 1

Biopanning was performed against eight non-BlCa cell lines, namely B16-PSMA, B16, VCaP, PC346C, LNCaP, MDAPCa2b, T47D and MCF7, and against six BlCa cell lines: JON, T24, J82, TCCSUP, RT112 and 5637.

Step 1: Ca cell type preparation

We used 1 x 10^7 cells for each cell line. For the cell preparation, we performed the following steps:

1. The cells were melted at room temperature.
2. We added RPMI-FBS 10% medium to dilute the cells 10-fold.
3. The cells were centrifuged at 1,200 rpm for 5 min (Eppendorf centrifuge 5804 R).
4. The supernatant was removed.
5. The pellet was resuspended in 5 mL 1% BSA (bovine serum albumin)-$_{1x}$PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 and 2 mM KH_2PO_4) and centrifuged again for 5 min at 1,200 rpm (Eppendorf centrifuge 5804 R).
6. The supernatant was removed again and the pellet was resuspended in 800 µL 1% BSA-$_{1x}$PBS.

Step 2: Mixing of the pre-selected libraries

Libraries were mixed at a ratio of 25:40:35 - preBlC-L1P4:preBlC-LUPCa1:preBlC-LUPCa2, where we aim for a total of 2.0 x 10^{11} phages (Table 3). The library concentrations were measured in 'Materials and Methods' section 3.2 step 9. We added 185 µL of 2% BSA to have a 1% BSA end concentration of the library mixture and incubated for ~ 15 min.

M13 libraries	Library concentration (phages/mL)	Ratio libraries	Number of phages added to the mixture	Volume (µL) used from library stock
PreBlC-L1P4	1.4 x 10^{12}	25	5 x 10^{10}	36
PreBlC-LUPCa1	6.9 x 10^{11}	40	8 x 10^{10}	116
PreBlC-LUPCa2	2.1 x 10^{12}	35	7 x 10^{10}	33
Number of phages in end volume			2.0 x 10^{11}	185 µL
2% BSA				185 µL
End volume preBlC-library				370 µL

Table 3 | Calculations for mixing the pre-BlCa selected libraries at a 25:40:35 ratio.

Step 3: Mixing of phages with Ca cells

We incubated each cell line with the library mixture (800 µL suspended cells + 370 µL of the library mixture) at room temperature for ~ 60 min while being gently mixed on a Heidolph duomax 1030. After incubation, they were centrifuged at 1,600 rpm (Eppendorf centrifuge 5417 R) for 5 min.

Step 4: TG1 culture ('Materials and Methods' section 3.2 step 4)

Step 5: Removal of unbound phages ('Materials and Methods' section 3.2 step 5)

Instead of 30 mL 0.5% BSA-$_{1X}$PBS for washing the cells, we used 10 mL.

Step 6: Elution ('Materials and Methods' section 3.2 step 6)

As we have used 1 x 10^7 cells, we used 300 µL H_2O / 200 mM trimethylamine / 1 M tris-HCl for this step, leading to a total volume of approximately 1.1 mL.

Step 7: Re-infection ('Materials and Methods' section 3.2 step 7)

As we ended with 1.1 mL after step 6, we added 10 mL TG1 culture to the phage supernatant (10:1 ratio). We incubated overnight in 300 mL TAG per cell line while being gently shaken at 270 rpm (New Brunswick Scientific C25 Incubated Floor Model Shaker).

Step 8: Titration See 'Materials and Methods' section 3.2 step 8.

Step 9: Next steps See 'Materials and Methods' section 3.2 step 9.

Step 10: Glycerol stocks See 'Materials and Methods' section 3.2 step 10.

Step 11: phage production

Using the bacteria from step 7, we produced phages using the same protocol used in 'Materials and Methods' section 3.3.

3.4.2 Selection of Nb-phages – round two

During this project, we performed a second panning round using the phages selected from the first panning. For this, we repeated the protocol for the first panning round (see 'Materials and Methods' section 3.4.1).

3.5 DNA isolation

We used the LabNed Plasmid kit to isolate plasmid DNA. The following protocol applies when using 25 mL of bacteria culture. The following protocol was adapted from LabNed *(2017)* [50].

Step 1: Equilibrating the column

The LabNed Midi Columns were placed on the rack and 10 mL equilibrium buffer (E4) was added consisting of 0.1 M Sodium Acetate (pH 5.0), 0.6 M NaCl and 0.15% (v/v) Triton X-100. We let the buffer run through the column by gravity flow.

Step 2: Preparing Cell Lysate

We harvested the bacteria by centrifuging at 4,000 rpm for 10 min (Beckman centrifuge J-6B). Secondly, we removed all supernatant and added 4 mL cell resuspending buffer (E1) with RNase A consisting of 50 mM Tris-HCl (pH 8.0) and 10 mM EDTA. The mixture was gently mixed until homogenous. Afterwards, we transferred the bacteria suspension to 15 mL centrifuge tubes. We added 4 mL lysis buffer (E2) consisting of 0.2 M NaOH and 1% (w/v) SDS, mixed gently until a homogenous mixture was obtained (Do not vortex!) and incubated for 5 min. Next, we added 4 mL precipitation buffer (E3) consisting of 3.1 M Potassium acetate (pH 5.0) and mixed by immediately inverting the tube. Finally, the mixture was centrifuged at 4,000 rpm for 30 min at room temperature (Beckman centrifuge J-6B).

Step 3: Binding and washing of DNA

The supernatant from *step 2* was loaded into the equilibrated column from *step 1*. The solution must flow through the column by gravity flow. Afterwards, the column was washed twice with 10 mL wash buffer (E5) consisting of 0.1 M Sodium Acetate (pH 5.0) and 0.8 M NaCl.

Step 4: Eluting and precipitating DNA

We placed a sterile tube under the column and 5 mL elution buffer (E6) consisting of 0.1 M Tris-HCl (pH 8.5) and 1.25 M (NaCl) was added to the column. The elution flows through the column into the sterile tube, which now contained the purified DNA. Next, 3.5 mL isopropanol (2-propanol) was added to the eluate and mixed well. Followed by centrifugation at 5,000 rpm for 30 min at 4ºC (Eppendorf

centrifuge 5804R). Next, the DNA pellet was washed in 1 mL 70% ethanol and again centrifuged at 7000 rpm for 5 min at 4ºC (Eppendorf centrifuge 5417R). The supernatant was carefully removed and discarded. Finally, the DNA pellet was air-dried for 10 min, diluted in 100 µL H_2O and stored at -20ºC.

Step 5: DNA concentration Nanodrop

Using the Thermofisher scientific Nanodrop ND-1000, we measured the DNA concentration. First, we selected the 'nucleic acid' option, initialized the instrument with water and added 1 µL H_2O as a reference measurement. Then, we added our DNA sample and the concentration is automatically calculated by the Nanodrop instrument in ng/µL.

3.6 Polymerase chain reaction

We performed a standard PCR with the isolated DNA samples from the selection assays. To each DNA sample, we added a specific forward and reverse barcode (Table 4), which can be observed in Supplementary Data Section 2. By using these specific barcodes, each sample can be distinguished during NGS. For PCR, we used primers containing both an Illumina-sequence needed for NGS readout and the barcode-sequence, needed to distinguish samples. We also added two negative controls to each PCR run: a primer without DNA and DNA without primers. In the end, we performed two NGS assays (NGS13 and NGS14) which included PCR samples 13.1-13.18 and 14.1-14.18.

PCR samples	DNA sample	Primer mixture	PCR samples	DNA sample	Primer mixture
13.1	preBIC-L1P4	B8	14.1	VCaP SR1 *	B8
13.2	preBIC-LUPCa1	C9	14.2	LNCaP SR1 *	C9
13.3	preBIC-LUPCa2	B12	14.3	PC346C SR1	B12
13.4	RT112 SR1 *	D7	14.4	MDAPCa2b SR1	D7
13.5	TCCSUP SR1 *	C10	14.5	T47D SR1	C10
13.6	T24 SR1 *	C12	14.6	MCF7 SR1	C12
13.7	JON SR1 *	B9	14.7	B16 SR1	B9
13.8	J82 SR1 *	D8	14.8	B16-PSMA SR1	D8
13.9	5637 SR1	D10	14.9	VCaP SR2 *	D10
13.10	RT112 SR2 *	H1	14.10	LNCaP SR2 *	H1
13.11	TCCSUP SR2 *	E6	14.11	PC346C SR2	E6
13.12	T24 SR2 *	F5	14.12	MDAPCa2b SR2	F5
13.13	JON SR2 *	G1	14.13	T47D SR2	G1
13.14	J82 SR2 *	G4	14.14	MCF7 SR2	G4
13.15	5637 SR2	G5	14.15	B16 SR2	G5
13.16	LUPCa1	E1	14.16	B16-PSMA SR2	E1
13.17	LUPCa2	H3	14.17	preBIC-L1P4	H3
13.18	L1P4	H4	14.18	L1P4	H4

Table 4 | DNA barcode combinations PCR. SR1 = selection round 1, SR2 = selection round 2. * No RNase was added during DNA isolation. All PCR samples 13.1-13.18 and 14.1-14.18 were included in the NGS runs.

3.6.1 PCR double-temperature run

For PCR, we used KAPA polymerase (Roche) at 0.02 U/µL final concentration, 0.3 mM dNTPs (dATP, dCTP, dGTP and dTTP), 1x HIFI (fid) buffer, 0.1 µM primer mix (forward + reverse) and 50 ng DNA. After adding the samples to a 96-well PCR plate, we performed a PCR double-temperature run using a Biometra thermocycler (Table 5).

Incubation time	Temperature	Times repeated
3 min	95 °C	
20 sec 20 sec 20 sec	98 °C 60 °C 72 °C	7x
20 sec 30 sec	98 °C 72 °C	20x
5 min	72 °C	
-	4 °C	

Table 5 | PCR double-temperature run cycle. Showing the different incubation times per temperature and how many times a certain cycle is repeated.

3.6.2 PCR clean-up

Adapted from Wizard(R) SV Gel and PCR Clean-Up System protocol [51]. We used the Micro Star 17 centrifuge for all steps.

Step 1: Processing PCR amplifications

We added membrane binding solution (MBS) (4.5 M guanidine isothiocyanate + 0.5 M potassium acetate (pH 5.0)) to the PCR amplification at an equal volume.

Step 2: Binding of DNA

The SV minicolumn was inserted into the collection tube. The PCR product was transferred to the minicolumn and we incubated at room temperature for 1 min. Next, we centrifuged at 12,000 rpm for 1 min. The flowthrough was discarded and the minicolumn was reinserted into the collection tube.

Step 3: Washing

We added 700 µL membrane wash solution (MWS) (80% ethanol, 10 mM potassium acetate (pH5.0), 16.7 µM EDTA (pH 8.0)) to the minicolumn and centrifuged at 14,000 × g for 1 min. Again, the flowthrough was discarded and the minicolumn was reinserted into the collection tube. This step was repeated with 500 µL MWS and centrifuged at 14,000 rpm for 5 min. Finally, the collection tube was emptied and re-centrifuged for 1 min at 14,000 rpm.

Step 4: Elution

We transferred the minicolumn to a clean 1.5 mL Eppendorf tube and added 50 µL of Nuclease-Free Water (or TE-buffer) to the minicolumn. Followed by incubation at room temperature for 1 min. Finally, the minicolumn was centrifuged at 14,000 rpm for 1 min. The minicolumn can now be discarded and the Eppendorf tube containing the PCR amplification product is stored at – 20°C.

3.7 Qubit dsDNA high-sensitivity assay

After PCR clean-up, the DNA concentration of the different samples was measured using the Qubit dsDNA high-sensitivity (HS) assay kit with the Qubit Fluorometer. Using the calculated concentration, we diluted the DNA concentration to 4 nM with 10 mM Tris-Cl, pH 8.5, before proceeding to NGS.

3.8 Next-generation sequencing

For both NGS13 and NGS14, we added 10 µL of the samples at 4 nM concentration together. Both NGS13 (including PCR samples 13.1-13.18) and NGS14 (including PCR samples 14.1-14.18) were performed by dr. Eric Bindels, department of Haematology. He used the Illumina MiSeq V3 2x300 bp kit, 7.5% phiX library, with 10 pM input.

3.9 Preparation of FroCellCA

During this project, we used a method of immunocytochemical staining to validate our libraries and phage-Nbs. For this validation method, we prepared dried cell arrays containing 21 different cell lines: six BlCa (JON, J82, T24, TCCSUP, RT112, 5637), seven PCa (VCaP, LNCaP, 22RV1, DU145, PC346C, PC3, MDAPCa2b), colorectal Ca COLO205, liver Ca Hep3B, primary Human Umbilical Vein Endothelial HUVEC, murine melanoma B16, breast Ca MCF7, lung Ca NCI-H460 and HEK293 (Figure 7). After cell culturing (see 'Materials and Methods' section 3.1), the cells were counted and centrifuged for 5 min at 1,000 rpm (Eppendorf centrifuge 5804 R). The supernatant was removed, the pellet was washed with PBS and afterwards resuspended in an appropriate volume of H_2O or PBS leading to 1 x 10^6 cells/mL. The cells were seeded onto glass microscope slides containing 21 wells, and 5 µL of cells were added per well leading to a total of 5,000 cells. The glass slides were air-dried overnight in a flow chamber at room temperature and afterwards stored at -20°C. The slides with seeded cells were named: Frozen Cell Cancer Arrays (FroCellCAs).

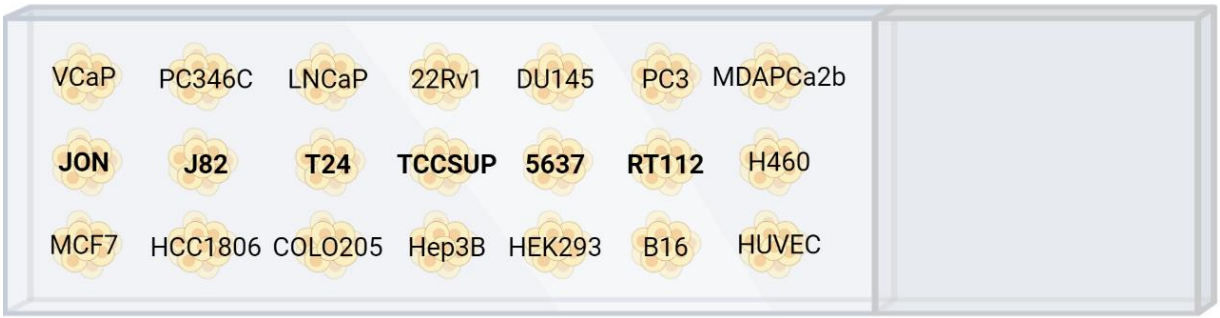


Figure 7 | FroCellCA cell line composition on a glass slide. BlCa cell lines are shown in Bold. *Created with BioRender.com.*

3.10 Immunocytochemical staining of FroCellCAs

Step 1: Preparing the slides

First, the FroCellCAs were placed in a sequenza rack and washed with 1 mL $_{1x}$PBS for 5 min, followed by blocking of the slides with 1 mL 1% BSA 3 x 5 min.

Step 2a: Phage incubation

We incubated the slides with 2 x 200 µL of the phages at 5 x 10^{10} phages/mL concentration for 30 min. Next, the slides were washed 3 x 5 min with 1 mL 0.2% BSA.

Step 2b: Antibody incubation

When using antibodies instead of phages, we incubated the slides with 2 x 200 µL anti-CD9-biotin or anti-CD63-biotin at 2.5 ng/µL for 30 min and washed by 3 x 5 min with 1 mL 0.2% BSA.

Step 3a: AntiM13-HRP incubation

AntiM13-HRP is used when slides are incubated with phages. We diluted antiM13-HRP (2 mg/mL) 250x in 1%BSA, added 2 x 200 µL to the slides and incubated for 30 min, followed by 3 x 5 min washing with 1 mL 0.2% BSA.

Step 3b: streptavidin-HRP incubation

Streptavidin-HRP is used in the case of biotin-labelled antibodies. The initial concentration of 0.82 g/L streptavidin-HRP was diluted 500x. We added 2 x 200 µL to the slides and incubated for 30 min, followed by 3 x 5 min washing with 1 mL 0.2% BSA.

Step 4: DAB staining

The bound antiM13-HRP / streptavidin-HRP is stained with peroxidase/DAB using the Dako REAL EnVision Detection system. The peroxidase/DAB was mixed with the substrate buffer at a 1:50 ratio. We added 2 x 200 µL/slide for 5 min and washed the slides with 1 mL 0.2% BSA 3 x 5 min.

Step 5: Imaging

For imaging, we used the BX41TF Olympus microscope. The slides were analysed using Cell B software with a manual setting and 100-200 ms conversion time. Afterwards, the images were further analysed using ImageJ. Within one experiment, it is important to keep imaging settings consistent as differences in these settings lead to differences when analysing the staining strength.

3.11 Hematoxylin-Eosin staining of FroCellCAs

The FroCellCAs were stained with Hematoxylin-Eosin to estimate the number of cells per well. First, we stained with hematoxylin for 2 min, followed by washing with water for 10 min. Secondly, we stained with eosin for 30 sec, again followed by washing with water for 5 min. The FroCellCAs were dried and imaged using the BX41TF Olympus microscope.

3.12 Gel electrophoresis

For gel electrophoresis, we used a 1% agarose gel containing 1x TRIS-borate-EDTA and stained the gel with 1 µL EtBr (10 mg/mL) / 100 mL gel. We loaded DNA samples (5-6 µL) in the gel box using an orange DNA loading dye 6x (2 µL) and ran the gel for 60-120 min at 150 V. The gel was imaged using the Isogen Imaging System ProXima 10 Phi.

3.13 Transformation of *E. coli*

For transformation of the *E. coli*, we used heat shock-competent bacteria which were prepared by Joke Veldhoven-Zweistra using the Inoue method [52]. The competent bacteria were placed on ice and we added 100-200 pg phagemid DNA to the bacteria. Important to keep in mind that the DNA volume may not exceed 5% of the volume of the competent cells. The phagemid DNA and bacteria were stored for 30 min on ice, followed by a 50 sec heat shock in a pre-heated water bath of 42°C, during which the plasmid DNA was transformed into the bacteria. Immediately after, the tubes were transferred to an ice bath and cooled for 2 min. Next, 900 µL of 2xTY medium was added and the tubes were incubated for 45 min at 37°C while gently shaking (160 rpm) (New Brunswick Scientific C25 Incubated Floor Model Shaker). Finally, 200 µL bacteria were plated on a 2xTY agar plate containing ampicillin and were grown overnight at 37°C (New Brunswick Scientific C25 Incubated Floor Model Shaker). The next day, a single colony was picked from the plates using a pipette tip and was grown in TAG for 16 hours. These infected bacteria were used for phage production (see 'Materials and Methods' section 3.3).

3.14 Restriction Enzyme Digestion of Nb vectors

To validate whether our Nb sequence was correctly inserted into the Nb-phage vectors ordered by Twist, we performed a restriction enzyme assay. We cut our Nb vectors using restriction enzymes NotI and EcoRI (both ordered from Roche) to retrieve two dsDNA strands of lengths 3,579 and 1,295 bp (Figure 8). We added 1 µg of the DNA vectors to 1 U enzyme, diluted 10x buffer H (50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl) to 1x from Roche and diluted to 50 µL using H_2O. The mixture was mixed gently by pipetting and restriction was performed at 37°C for 1 hour. After restriction, we performed gel electrophoreses and loaded 10 µL of the restriction mixture (see 'Materials & Methods' section 3.12).

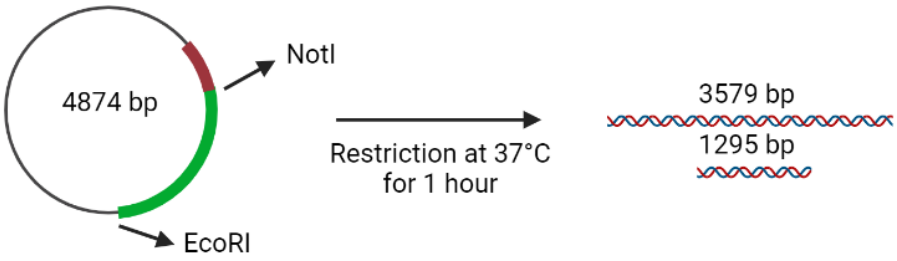


Figure 8 | Schematic overview of the restriction of our Nb-phage vectors with enzymes NotI and EcoRI. The Nb-phage vector has a length of 4,875 bp, where NotI cuts at bp 501 and EcoRI cuts at bp 1,663, leading to two strands of dsDNA of lengths 3,579 and 1,295 bp (green part in vector). The Nb sequence is depicted in red and is in front of the NotI restriction site. Restriction is performed at 37°C for 1 hour. *Created with BioRender.com.*

3.15 Enzyme-Linked Cell Assay

Step 1: Addition of cells/tissue to the 96-well plate

We performed a whole cell/meshed tissue enzyme-linked cell assay (ELCA). First, the cells and meshed tissues were suspended in EDTA and were spun for 5 min at 2,000 rpm (Eppendorf centrifuge 5417 R). The supernatant was removed and the pellet of cells was resuspended in 1 ml of 2% BSA-$_{1x}$PBS. We added two different counts for each cell line in 50 µL in duplicates to a V-shaped-bottom 96-well plate. For the experiments, we used either 1.5 x 10^{10} and 3.0 x 10^{10} cells or 3.0 x 10^{4} and 6.0 x 10^{4} cells. For the meshed tissues, we resuspended them in 500 µL 2% BSA-$_{1x}$PBS and dr. Raheleh Tooyserkani prepared 5x and 10x dilutions of the tissue samples. We added the tissue samples in 50 µL volume in duplicates to the plates. Also, we included two empty wells to test for unspecific binding of phages to the plates.

Step 2: Phage incubation

We used H6 as a positive control phage, NoBi as a negative control phage, and A7 as a PSMA-specific control phage and included phages BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 for this assay. All phages were diluted in 2% BSA-$_{1x}$PBS to a final concentration of 5 x 10^{10} phages/mL. We incubated the cells/tissues on the 96-well plates with 50 µL phages for 60 min while being shaked at 600 rpm on an IKA MTS 2/4 digital Microplate shaker.

Step 3: Washing the plates

After phage incubation, the 96-well plates containing a mixture of phage-tissue/cells was spun for 5 min at 1,000 rpm (Eppendorf centrifuge 5417 R). The supernatant was discarded and 150 µL 2% BSA-$_{1x}$PBS was added. Again, the plate was spun for 5 min at 1,000 rpm (Eppendorf centrifuge 5417 R). The washing steps were repeated three times for each plate and finally, the supernatant was removed.

Step 4: Incubation with antiM13-HRP

We diluted antiM13-HRP (2 mg/mL) 2,000x in 2%BSA, added 100 µL to each well on the plates and incubated for 60 min while being shaked at 600 rpm on an IKA MTS 2/4 digital Microplate shaker.

Step 5: Washing the plates

Repeat *step 3*.

Step 6: Addition of the OPD-substrate

For 10 mL OPD-substrate, we dissolved one tablet OPD (10 mg) in 5 mL H_2O, 2.57 mL Na_2PO_4 (0.2 M), 2.43 mL citric acid (0.1 M) and 5 µL H_2O_2 (30%). Next, we added 100 µL OPD-substrate to each well on the plates and incubated for 10 min.

Step 7: Plate reading

The plates were measured with the BIO-RAD microplate reader (Model 550) and the yellow-orange product caused by the OPD was read at 450 nm (Figure 9).

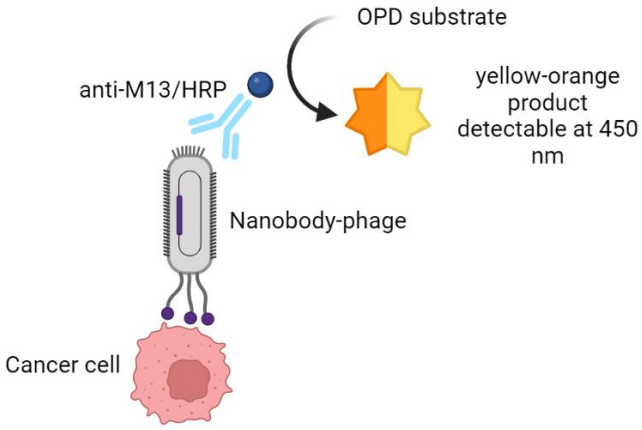


Figure 9 | Schematic overview of cell/tissue ELCA. Within the well, a Nb-phage binds to the antigens presented on the outer surface of a Ca cell. These Nb-phages are recognized by anti-M13/HRP. OPD is converted into a yellow-orange product by HRP and is detected at 450 nm. *Created with BioRender.com.*

3.16 Data analysis

3.16.1 FroDA score

Bright-field images obtained from the FroCellCA slides with the confocal microscope were processed in ImageJ using the code shown in Supplementary Data section 3. Using this macro, the average red, green, and blue (RGB) grey scale values of the DAB images were obtained. Using this code, we calculated a score, which we named the FroDA score, corresponding to specific RGB colours using the formula depicted in Figure 10. Higher scores correspond to more DAB staining, which is a measure of the number of bound Nbs to a specific cell line. The scores were corrected with the score of the Nobi Nb staining against each cell line, leading to a final score named: FroDA (FroCellCA Dab Analyses) score. For each experiment, the standard deviations were calculated together for all Nbs, leading to similar equations for each assay (Figure 10).

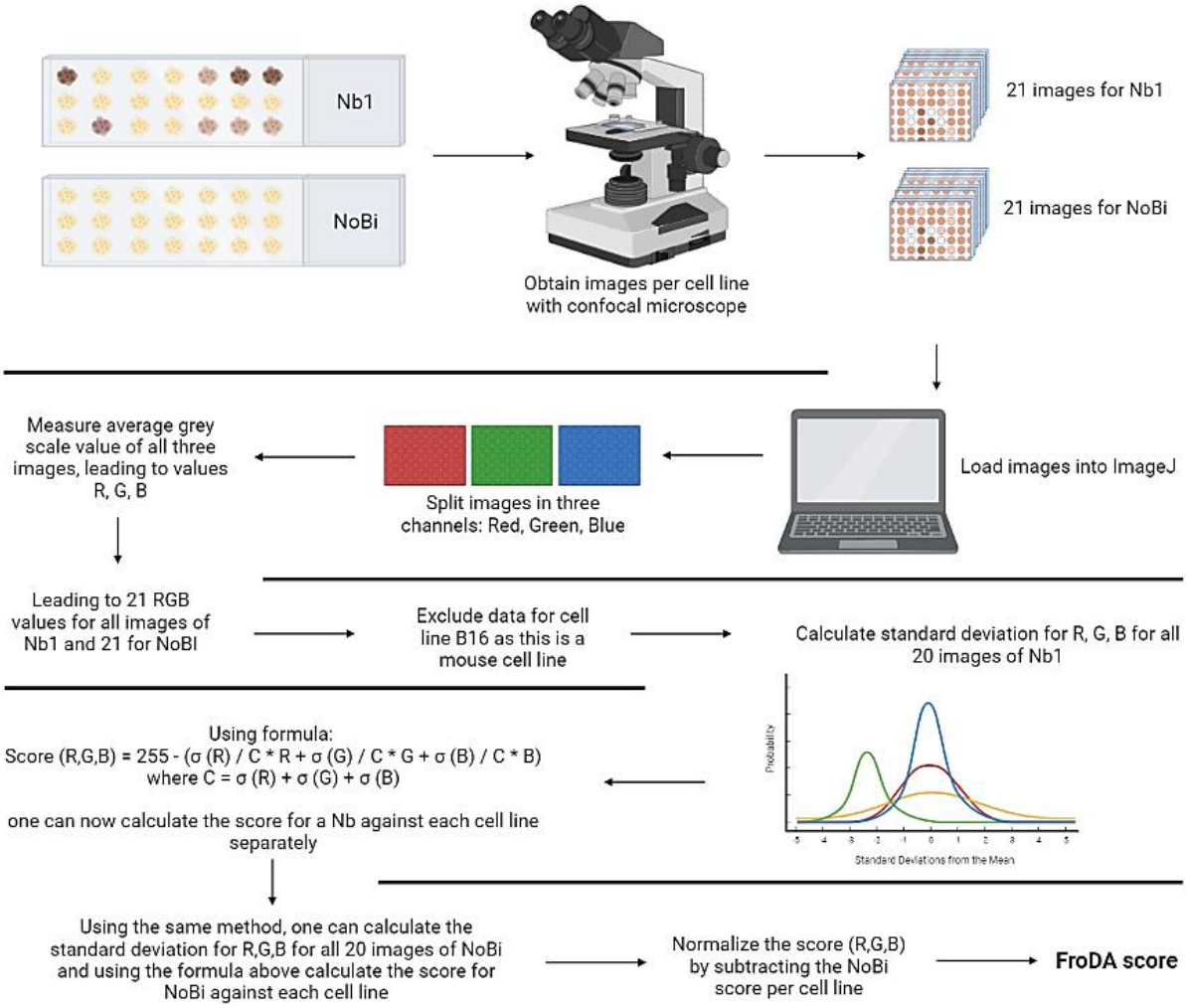


Figure 10 | Schematic overview of FroDA score calculation. A step-wise overview going from FroCellCA slides stained with imaginary Nb1 and NoBi to the FroDA score per Nb per cell line. *Created with BioRender.com.*

3.16.2 NGS analyses

A combinational script making use of PEAR, CutAdapt, MiXCR, BASH, R-script Single amino acid Difference (SaaDiff) and Excel was used to process the NGS output files containing the forward and reverse strand sequences. Firstly, the raw forward and reverse Illumina reads were paired using the PEAR software version 0.9.6 [53] within a snakemake data analysis workflow using the language of python. Within this workflow, also FastQC, CutAdapt and MiXCR were executed. FastQC performs a quality control of the sequence data [54]. CutAdapt was used to match sequences to a specific sample based on forward and reverse primer combinations [55] and MiXCR matches the sequences to known VHH-sequences from the 'Alpaca' species to retrieve the CDR3 sequence [56]. Secondly, a script written in BASH was used to count the number of specific CDR3 sequences for each sample. Thirdly, the R-script SaaDiff was used to clean up the CDR3 counts, remove the first and last AA, make a cut-off at 300,000 CDR3 sequences and generate an Excel file containing the top 300,000 CDR3 sequences. Also, a family tree of these sequences was made, where family 1 contains the most counts and where families 1 and 2 only have one amino acid in difference, families 1 and 3 have two AAs in difference,

and so on. For more information on family tree construction see ‘Results’ section 4.3. Finally, both runs for NGS13 and NGS14 are combined into one final NGS database containing the top 300,00 CDR3 sequences for each sample. Scripts were written by Rick Jansen, MSc and were executed together with prof. dr. Guido Jenster.

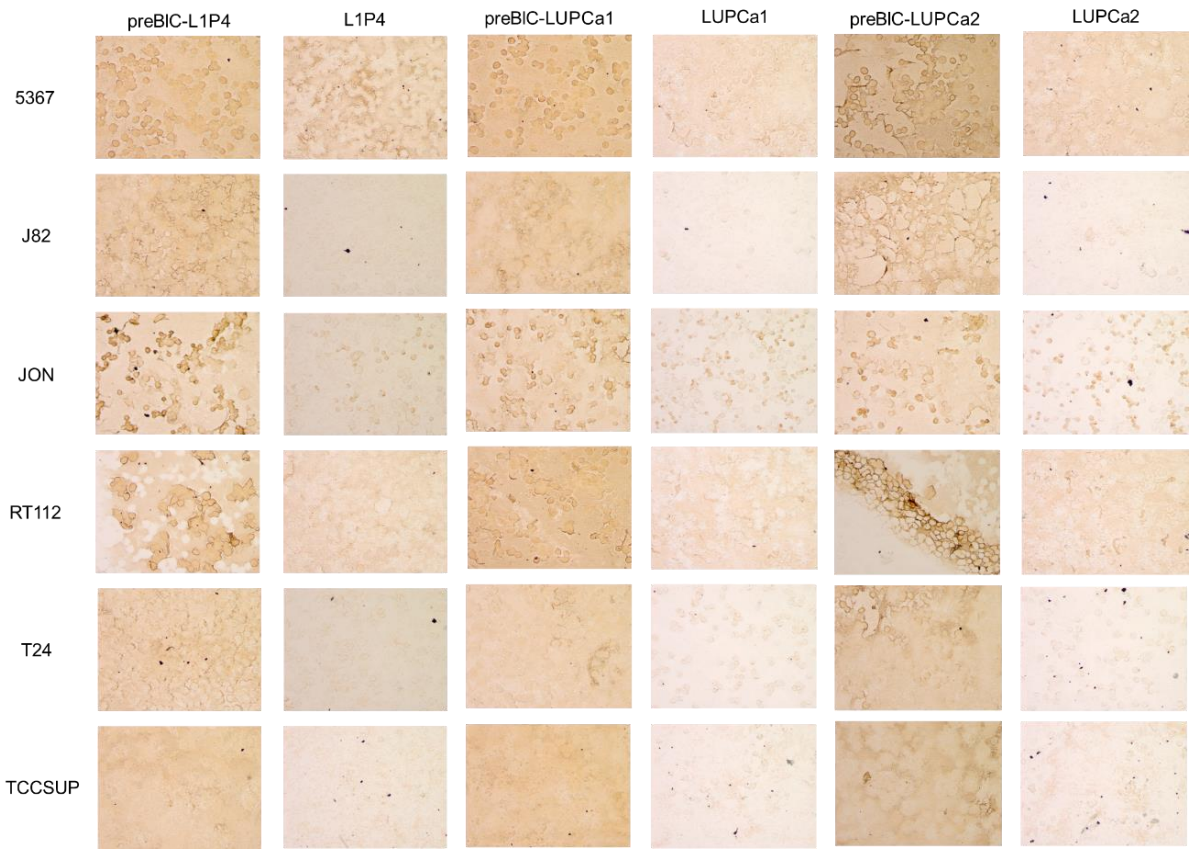
4 Results

4.1 *In vitro* immunocytochemical staining showed that pre-selected Nb libraries have an increased affinity towards Ca cell lines

Previously, three M13-phage Nb libraries (L1P4, LUPCa1 and LUPCa2) have been used for biopanning selection of Ca-specific Nbs. For this project, we have pre-selected these three libraries on six BlCa cell lines (RT112, TCCSUP, T24, JON, J82, 5637) and performed immunocytochemical staining using FroCellCAs to check the effect of pre-selection (see 'Materials and Methods' section 3.9 & 3.10). In total, we stained nine FroCellCAs, three with our pre-selected libraries (preBlC-L1P4, preBlC-LUPCa1 and preBlC-LUPCa2), three with the original Nb libraries (L1P4, LUPCa1, LUPCa2) and three with control-group Nbs, namely H6, A7 and NoBi. All FroCellCAs were stained twice with 200 µL Nbs at 5 x 10^{10} phages/mL concentration, followed by incubation with HRP-labelled anti-M13 Abs after which DAB staining was added to visualize the signal strength of Nb-phage binding to each cell line.

First, we will elaborate on the used control Nbs: H6, NoBi and A7. H6 binds CD9, which is present on almost all cell types. For H6, we indeed observed high signal strengths for all cell lines on the FroCellCa, except for Hep3B and HUVEC (Supplementary data, section 4: Figure 34). NoBi is a non-binding Nb and, as expected, we observed no DAB staining signal for all cell lines (Supplementary data, section 4: Figure 35). Lastly, we used A7 as a PSMA binding Nb, leading to specific staining of PSMA-positive cell lines. On our FroCellCAs, we have included four PSMA-positive PCa cell lines, namely LNCaP, PC346C, 22Rv1 and MDAPCa2b [57]. Indeed, we observed specific binding of A7 to these cell lines (Supplementary data, section 4: Figure 36). To make the analyses more quantitative, we calculated the FroDA scores (see 'Materials and Methods' section 3.16) for the FroCellCa incubated with A7. Comparing the FroDA scores between PSMA and non-PSMA-expressing cell lines, we observed higher FroDA scores for PSMA-expressing cell lines (FroDA LNCaP = 47, PC346c = 57, 22RV1 = 51, MDAPCa2b = 68) compared to non-PSMA-expressing cells ($\mu = 27$, $\sigma = 3.6$). From these controls, we conclude that the experimental set-up of the FroCellCA worked and provides binding patterns and objective FroDA scores as expected.

We observed a stronger signal of DAB for the slides stained with pre-selected libraries (Figure 11) compared to the previously used parental Nb libraries (Supplementary data, section 5: Figure 37 – 39). To further analyse the FroCellCAs, we calculated the FroDA scores and observed higher scores for slides stained with pre-selected libraries. As the libraries were pre-selected on BlCa cell lines, we initially expected increased FroDA scores for only BlCa cell lines, however, this stronger staining was also observed for other Ca cell lines. We hypothesize that pre-selection serves as a general selection, mostly for common cell membrane proteins. Whether BlCa-specific Nbs present in the pre-selected libraries show stronger binding, will become clear upon individual cell line pannings and database comparison.



preBIC-L1P4			L1P4	preBIC-LUPCa1			LUPCa1	preBIC-LUPCa2			LUPCa2
5367	34	→	19	**5367**	36	→	0	**5367**	51	→	6
J82	49	→	31	**J82**	44	→	-4	**J82**	39	→	-1
RT112	24	→	14	**RT112**	25	→	-1	**RT112**	18	→	-9
JON	36	→	28	**JON**	49	→	7	**JON**	45	→	13
T24	39	→	30	**T24**	33	→	-3	**T24**	45	→	-1
TCCSUP	38	→	10	**TCCSUP**	44	→	-4	**TCCSUP**	50	→	1
22Rv1	36	→	9	22Rv1	36	→	-7	22Rv1	42	→	-5
COLO205	29	→	31	COLO205	25	→	-2	COLO205	36	→	0
DU145	40	→	14	DU145	36	→	-2	DU145	84	→	-3
H460	33	→	13	H460	43	→	-2	H460	37	→	0
HCC1860	24	→	28	HCC1860	18	→	-8	HCC1860	15	→	-4
HEK293	40	→	22	HEK293	44	→	-2	HEK293	61	→	2
Hep3B	39	→	10	Hep3B	40	→	-4	Hep3B	52	→	-1
HUVEC	15	→	14	HUVEC	24	→	1	HUVEC	29	→	8
LNCaP	35	→	9	LNCaP	44	→	1	LNCaP	40	→	-1
MCF7	40	→	34	MCF7	36	→	7	MCF7	30	→	0
MDPCa2b	35	→	13	MDPCa2b	48	→	-3	MDPCa2b	53	→	1
PC3	34	→	13	PC3	48	→	-5	PC3	24	→	0
PC346C	40	→	34	PC346C	41	→	-3	PC346C	27	→	0
VCaP	32	→	32	VCaP	35	→	24	VCaP	30	→	-1

Figure 11 | preBlC-library versus parental library binding to Ca cell lines. (Top) Images of DAB signal strength of six BlCa cell lines (5637, J82, JON, RT112, T24, TCCSUP) stained with preBlC-library and parental library. (Bottom) The FroDA score is calculated for all cell lines comparing preBlC-libraries to the parental libraries. (Left) preBlC-L1P4 versus L1P4 binding. (Middle) preBlC-LUPCa1 versus LUPCa1 binding. (Right) preBlC-LUPCa2 versus LUPCa2 binding.

4.2 Nb counts within control groups L1P4 and preBlC-L1P4 are linearly correlated when generated in two separate NGS databases

Using the pre-selected libraries, we performed two consecutive selection (panning) rounds against JON, J82, TCCSUP, T24, RT112, 5637, VCaP, LNCaP, PC346C, MDAPCa2b, T47D, MCF7, B16 and B16-PSMA. After each selection, the DNA from the Nbs was isolated and amplified using PCR. The PCR products were used to generate an NGS database containing the binding levels (BLs) of the top 300,000 Nbs against different cell lines. A final NGS database was generated by combining the BLs of Nbs found within NGS13 and NGS14. As a quality control for reproducibility of NGS, data analysis and database construction, we added control groups L1P4 and preBlC-L1P4 to both NGS13 and NGS14.

We plotted the counts (number of hits) for the top 300,000 Nbs found within the L1P4 sample of NGS13 against NGS14 and observed a linear trend with $R^2 = 0.9983$ (Figure 12). Similarly, we observed a linear trend for the Nbs found within preBlC-L1P4 between NGS13 and NGS14, with $R^2 = 0.9994$ (Figure 12). This shows that the NGS databases from NGS13 and NGS14 can be combined into one final database without the need for batch correction.

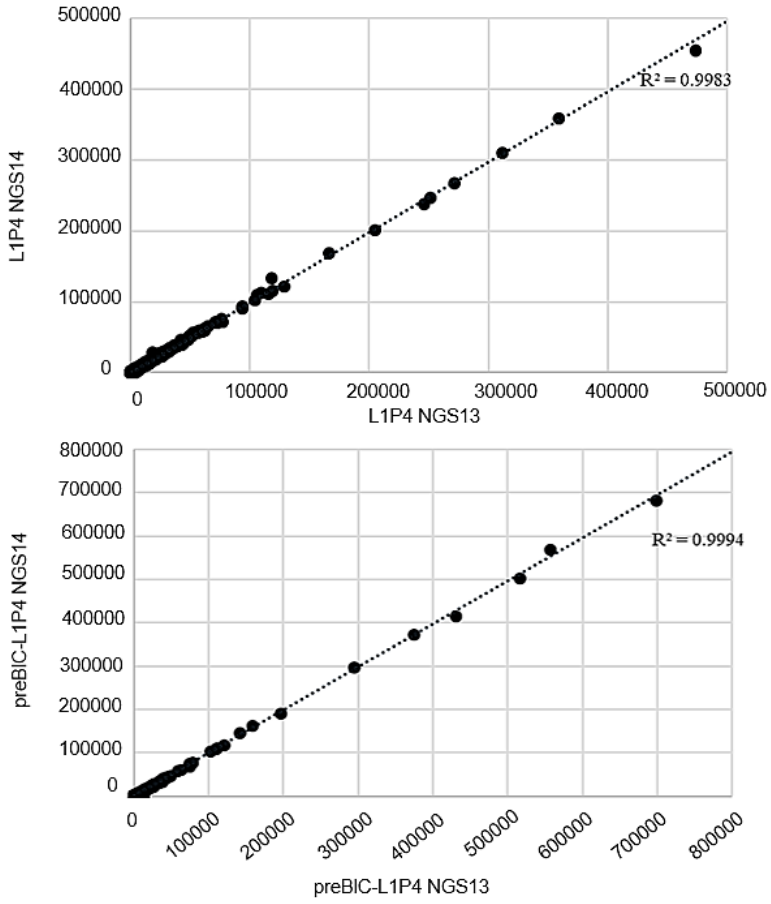


Figure 12 | Top 300,000 Nbs and their counts within L1P4 and preBlC-L1P4 comparing NGS13 to NGS14. (Top) The counts for the top 300,000 Nbs found within sample L1P4 in NGS13 (x-axis) plotted against L1P4 within NGS14 (y-axis). (Bottom) We plotted the counts for the top 300,000 Nbs found within sample preBlC-L1P4 in NGS13 (x-axis) against preBlC-L1P4 in NGS14 (y-axis).

4.3 Pre-selection of M13-phage libraries on six BlCa cell lines is beneficial when searching for BlCa-specific Nbs

Nbs are sequenced based on their whole CDR1, CDR2 and CDR3 domains and when generating the NGS database, Nbs with identical CDR3 amino acid sequences are combined. Within this database, the BL of a particular CDR3 sequence against each cell line is shown. Therefore, every row of the 300,000 rows has a unique CDR3 protein sequence, but the counts in the database consist of a compilation of counts of 'CDR3-identical' Nbs, typically with differences in CDR1 or CDR2. Although each row consists of a compilation of Nbs with identical CDR3, but potential differences in CDRs, we refer to the unique CDR3 sequences as a 'Nb'.

It happens frequently that a CDR3 sequence is very similar to another CDR3 sequence in another row. Using the SaaDiff script, all CDR3s with the same length and one amino acid difference are grouped into a CDR3 family. This family connection, the length of the CDR3 sequence and the CDR3 sequence itself is shown for each row.

To dive a little bit deeper into the family tree construction, we will elaborate on an example Nb tree. The family member order is depicted as @X@X@X@.., where the first X in **bold** indicates the family member indicator of that particular Nb and the lowest X in red is the family parent. In Figure 13, one can observe this example Nb family tree. The Nb with the highest total BL (sum of BLs against all cell lines) has CDR3 sequence 'CCCCC' (**@1**@2@7) and is named the parent of this tree. The family member information is depicted as **@1**@2@7, meaning that parent @1 has two children with indicators @2 and @7. The indicator order is based on total BL, where Nbs with higher total BLs have higher indicators than Nbs with lower total BLs. Nb @7 has amino acid sequence 'CDCCC' and has three children: 'DDCCC' (**@8**@7), 'CDDCC' (**@10**@7) and 'CDCDC' (**@9**@7). Nb @2 with 'CCECC' has one child with amino acid sequence 'CEECC' (**@3**@2@4@5@6), which itself has three children: **@4**@3 with 'CEEEC', **@5**@3 with 'CEECE' and **@6**@3 with 'EEECC'. We expect that the majority of family members recognize the same epitope due to their high similarity in CDR3 Nb sequence. During our Nb selection, we therefore take into account whether Nbs are part of a large family, whether they are the parent or a low abundant child, and whether the more abundant family members have the same binding pattern to the different cell lines. If so, we consider the selected Nb more reliable and typically select the most abundant Nb (parent) for further validation.

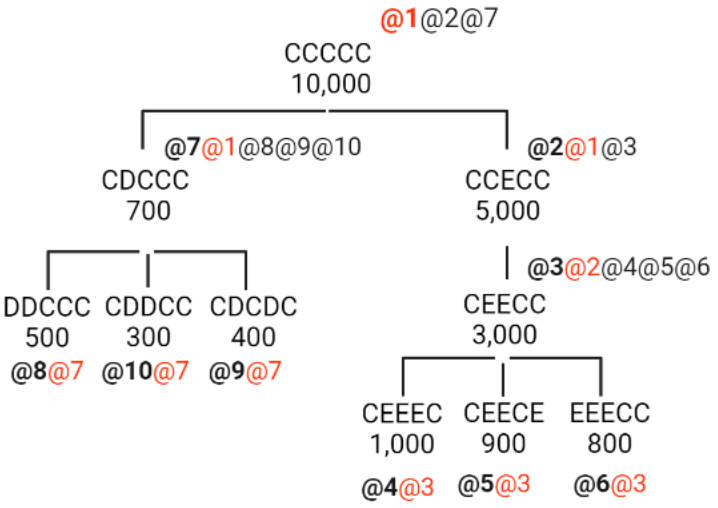


Figure 13 | Example family tree of Nbs. The sequence of each Nb is depicted, starting with family member 'CCCCC', which has children 'CDCCC' and 'CCECC', etc. The family member information is depicted as @X@X@X…, where the number in red is the family parent and the first number in **bold** is the family indicator of that Nb.

Using our NGS database, we searched for BlCa-specific Nbs using certain criteria. For this search, we used the following criteria:

- **Total BL** = sum of binding levels for one Nb against all Ca cell lines.
- **SR2/SR1$_{BlCa}$** = Ratio between the average (AVR) BL of Nbs to BlCa cell lines (T24, JON, J82, TCCSUP, RT112, 5637) after selections one (SR1) and two (SR2).
- **BlCa/Ca$_{SR1}$** = Ratio between the AVR BL of Nbs to BlCa cell lines and other Ca cell lines during SR1.
- **BlCa/Ca$_{SR2}$** = Ratio between the AVR BL of Nbs to BlCa cell lines and other Ca cell lines during SR2.
- **BlCa/BlN$_{tissue}$** = Ratio between the AVR BL of Nbs to BlCa tissue and Bladder Normal (BlN) tissue samples. This was data obtained from an NGS database generated in a previous project.
- **SR2/SR1$_{nonBlCa}$** = Ratio between the AVR BL of Nbs to non-BlCa cell lines (LNCaP, VCaP, MDAPCa2b, PC346C, B16, B16-PSMA, MCF7 and T47D) after SR1 and SR2.
- **SR2/SR1$_{BlCa/Ca}$** = Ratio between BlCa/Ca$_{SR2}$ and BlCa/Ca$_{SR1}$.
- **B16-PSMA/B16-WT=** Ratio between the BL of Nbs against B16-PSMA expressing cells and B16 wild type (WT) cells within the proprietary database.

We performed analyses on our newly generated NGS database to prove the value of a pre-selection prior to selection. First, we selected Nbs using the following criteria: total BL > 200, ratio BlCa/Ca$_{SR1}$ > 10 and ratio BlCa/Ca$_{SR2}$ > 20. This led to 214 Nbs out of 300,000, which are expected to behave in a BlCa-specific manner. Among these suspected 214 BlCa-specific Nbs, we observed that 60% of these Nbs originate from preBlC-LUPCa1, 23% from preBlC-LUPCa2 and 17% from preBlC-L1P4. As most Nbs

originate from preBlC-LUPCa1, we focused on this pre-selected library. We plotted the top 300,000 Nbs and their counts within preBlC-LUPCa1 against the top 300,000 Nbs and their counts within LUPCa1. We observed a linear distribution between the counts of Nbs within the two libraries (Figure 14 top). However, when plotting the top 214 Nbs specifically selected for BlCa cell lines, we observed that these Nbs have higher counts in the preBlC-LUPCa1 compared to LUPCa1 (Figure 14 bottom). This shows that the abundance of BlCa-specific Nbs is higher within the pre-selected library and pre-selection is therefore beneficial when searching for BlCa-specific Nbs.

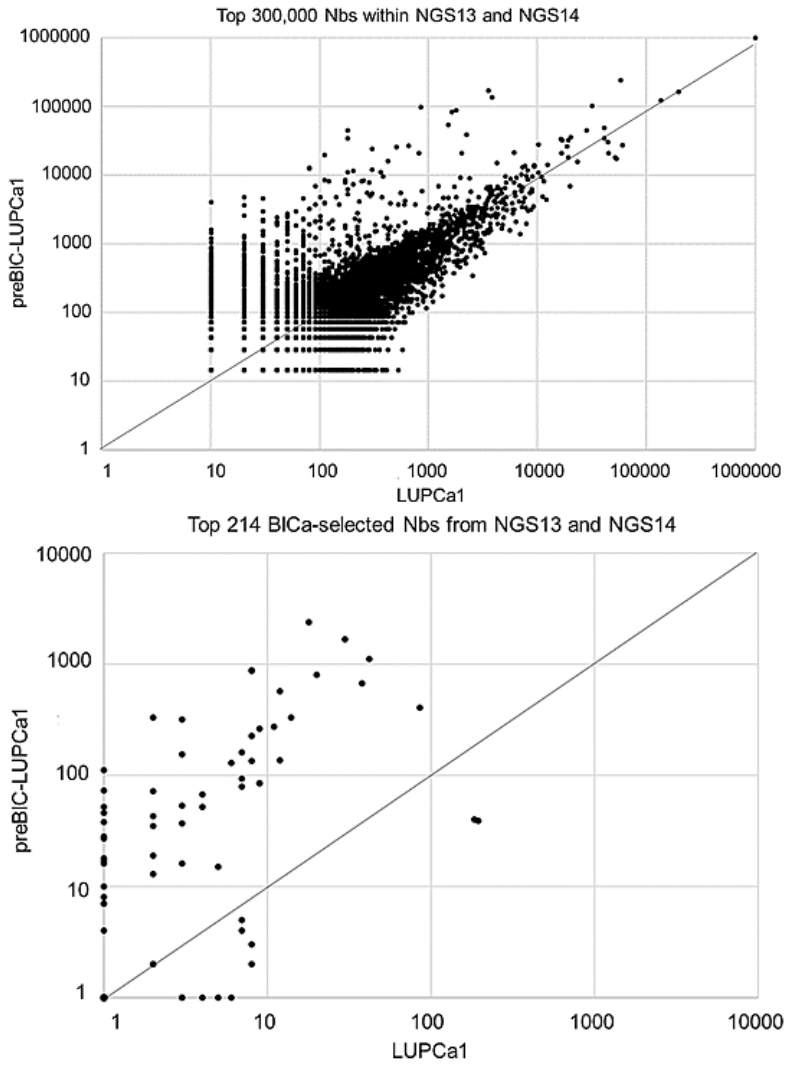


Figure 14 | Top 214 BlCa-specific Nbs from NGS13 and NGS14 have higher counts within preBlC-LUPCa1 compared to LUPCa1. (Top) The top 300,000 Nbs are shown with their number of counts within the preBlC-LUPCa1 library (y-axis) and the LUPCa1 library (x-axis). (Bottom) The number of counts for the top 214 BlCa-selected Nbs within preBlC-LUPCa1 (y-axis) and LUPCa1 (x-axis) are plotted. The black line is at $R^2 = 1.000$. All axes are logarithmically scaled.

4.4 Two consecutive rounds of preBlC-libraries against BlCa cell lines increases the chance of finding BlCa-specific Nbs

Within the following section, we will elaborate on several findings that prove the value of two rounds when searching for BlCa-specific Nbs.

4.4.1 The BL of non-specific BlCa Nbs against BlCa cell lines decreases when performing two consecutive rounds of selection

First, we analysed 15 BlCa-Nbs which were selected and tested by another master's student, Jay Yee. He selected these BlCa-Nbs from the proprietary database containing the top 300,000 Nbs and their BLs against Ca cell lines, Ca tissue and healthy tissue. For validation of these Nbs, he performed immunocytochemical staining using FroCellCAs and ELCA. These experiments showed no specific BlCa binding of these particular Nbs. Within our newly generated NGS database, we were able to retrieve 13 of these Nbs. We observed that the AVR ratio $SR2/SR1_{BlCa/Ca}$ for these Nbs was 0.69 with σ = 0.28 (N = 13), meaning that the BLs of these Nbs against BlCa cell lines decreased when performing two consecutive rounds of selection. This correlation between validated non-specific Nbs and the absence of enrichment in round 2, indicates a means of eliminating irrelevant, non-specific Nbs.

4.4.2 The BL of B16-PSMA-specific Nbs against B16-PSMA cells increases when performing two consecutive rounds of selection

For the following analyses, we used the proprietary database to select PSMA-binding Nbs. We used the following criteria to do so: total BL > 100 and B16-PSMA/B16-WT > 1.5. Using these criteria, we selected 8,928 Nbs from 300,000 which show a slight specificity towards B16-PSMA cells compared to B16-WT cells, *in silico*. From these 8,928 Nbs, we retrieved 182 Nbs within our NGS database, where we excluded Nbs with a total BL against samples B16-PMSA SR1+SR2 lower than 100. For these Nbs, we found an AVR $SR2/SR1_{B16\text{-}PSMA}$ of 1.37 (N = 182), indicating that the BL of PSMA-specific Nbs increased when performing two rounds.

We repeated this analysis, where we selected non-specific PSMA Nbs based on the following criteria: total BL > 100 and ratio B16-PSMA/B16-WT < 0.5. We found 64,753 Nbs which show a slight specificity towards B16-WT cells compared to B16-PSMA cells. We retrieved 53 Nbs within our NGS database, again excluding Nbs with a total BL against samples B16-PMSA SR1+SR2 lower than 100, and found an AVR $SR2/SR1_{B16\text{-}PSMA}$ of 0.74 (N = 53). Indicating, that the BL of Nbs non-specific towards B16-PSMA decreased after the second selection.

4.4.3 Randomly selected Nbs show a decrease in BL when performing two consecutive rounds of selection

Next, we compared the AVR ratio $SR2/SR1_{B16\text{-}PSMA}$ for the N = 182 PSMA-specific Nbs against 10x 100 randomly selected Nbs from the NGS database. For these randomly selected Nbs, we found a ratio

$SR2/SR1_{B16\text{-}PSMA}$ of 0.778; 0.773; 0.880; 0.672; 0.944; 0.673; 0.997; 0.703; 0.950 with $\rho < 0.05$ (t-test Welch), and 1.055 with $\rho > 0.05$ (t-test Welch). We thus observe a decrease in BL against B16-PSMA for randomly selected Nbs when performing two rounds, whereas, we observed an increase in BL against B16-PSMA for PSMA-specific Nbs, again confirming that two rounds are beneficial when searching for cell line/marker-specific Nbs.

4.4.4 Nbs with $SR2/SR1_{BlCa} > 2.0$ have higher BLs against BlCa cell lines compared to PCa cell lines during both SR1 and SR2

For the top 300,000 Nbs, we plotted the AVR BlCa BL after SR1 against AVR PCa BL after SR1 and observed a linear distribution (Figure 15). When selecting Nbs based on the criteria $SR2/SR1_{BlCa} > 2.0$, we observe a shift of Nb BLs towards BlCa, while for $SR2/SR1_{BlCa} < 0.5$ we observe a shift of Nb BLs towards PCa (Figure 15 top). Similar results can be observed for SR2 (Figure 15 bottom). Again, these results show the additive value of performing two selection rounds, as this can be used as a criteria for the selection of Nbs, where they should have SR2/SR1 ratios > 1.0.

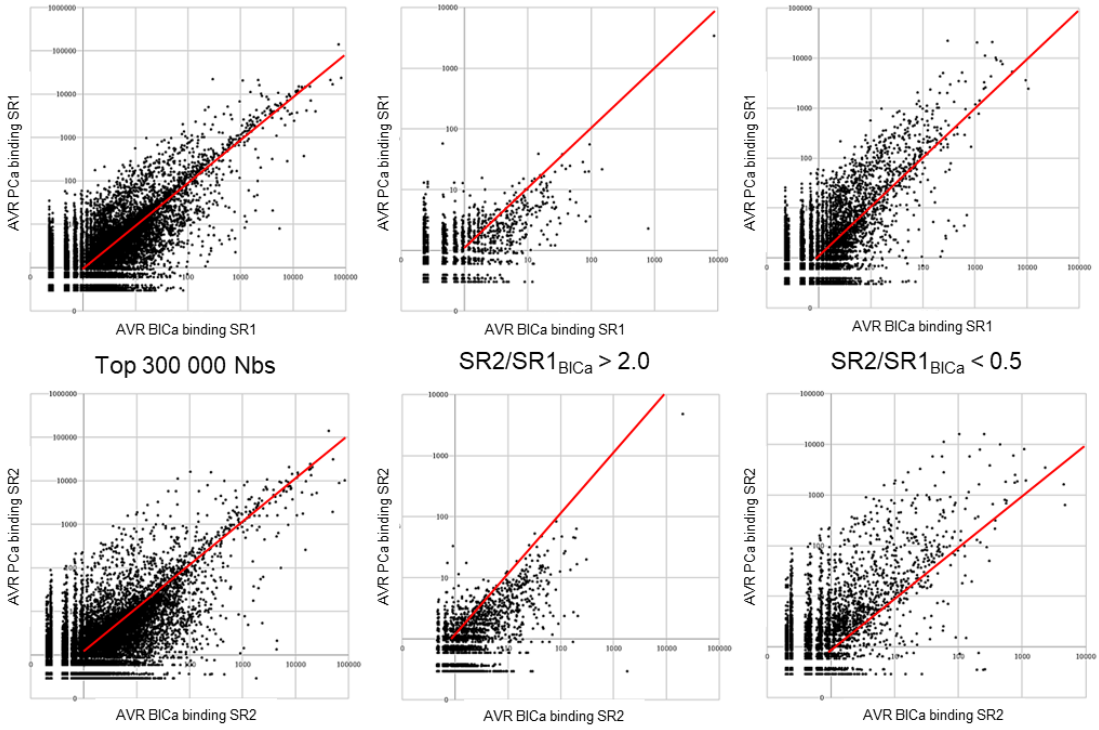


Figure 15 | Plots showing the AVR PCa BLs during SR1 (top) and SR2 (bottom) versus the AVR BlCa BLs. (left) All 300,000 Nbs present in the NGS database, (middle) Nbs that meet criteria $SR2/SR1_{BlCa} > 2.0$ and (right) Nbs that meet criteria $SR2/SR1_{BlCa} < 0.5$. The red line is where Nbs have equal AVR PCa (y-axis) and BlCa (x-axis) binding. All axes are logarithmically scaled.

4.4.5 During SR2 the BL of cell line specific Nbs become less dispersed

For all Nbs within our database, we calculated the AVR BL ratio BlCa/PCa for Nbs during SR1 and plotted this against the AVR SR2/SR1 BlCa ratio (Figure 16 top). We expected Nbs with higher BL in BlCa as compared to PCa (upper half of the plot) to be correlated to enrichment in SR2 against BlCa

cell lines (right part of the plot). Therefore, a cloud of Nbs in the upper right and bottom left of the plot. We observe a highly dispersed cloud of Nbs, with no clear relation between the AVR BL ratio BlCa/PCa during SR1 and the AVR SR2/SR1 BlCa ratio. However, when plotting the AVR BL ratio BlCa/PCa during SR2 of our Nbs, we do observe an increase in the expected relation with the AVR SR2/SR1 BlCa ratio. Nbs with a high SR2/SR1 BlCa ratio, also show high AVR BL BlCa/PCa ratios during SR2, while Nbs with low SR2/SR1 BlCa ratios, show low AVR BL BlCa/PCa ratios during SR2. From this, we conclude that the binding patterns of Nbs become more selective during SR2 and show a more cell line specific pattern.

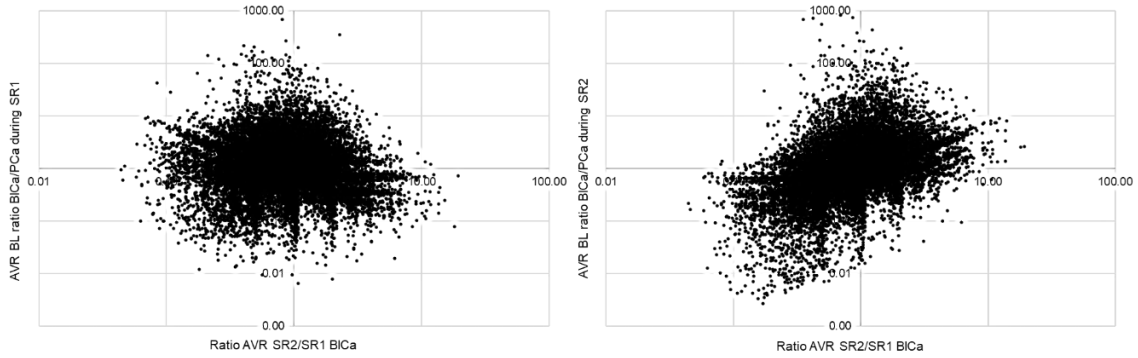


Figure 16 | BlCa/PCa ratios during SR1 (left) and SR2 (right) plotted against the ratio AVR SR2/SR1 BlCa. Both axes are logarithmically scaled. On the x-axis the ratio AVR SR2/SR1 BlCa is depicted and on the y-axis the AVR BL BlCa/PCa binding during SR1 (left) and SR2 (right) is depicted.

4.5 *In silico* selection of Nbs against BlCa

From our NGS database, we selected two non-specific and 10 BlCa-specific Nbs which were tested to confirm their specificity. We started with a global selection where we excluded Nbs that have a BL of zero for more than four BlCa cell lines and/or have more than one unknown amino acid or stop codon in their CDR3 sequence. After the global selection, we applied three selection criteria to the database (Figure 17):

Selection criterium 1: The total BLs of the Nb must be higher than 200, the AVR BL to BlCa cell lines must be at least 10-fold higher than the AVR BL to non-BlCa cell lines during SR1 and 20-fold higher during SR2. Also, the AVR BL against BlCa cell lines must be 1.5-fold higher during SR2 compared to SR1. Applying these criteria, we retrieved 49 Nbs from the 300,000 Nbs present in the NGS database.

Black and white screening: Secondly, we excluded Nbs with BLs higher than zero for non-BlCa cell lines. This led to 14 Nbs. Further selection was based on the hypothesis that Nbs from the same family follow the same binding trend against Ca cell lines. Using *family screening*, we were able to select two Nbs (**BlCaNb35** and **BlCaNb36**) originating from family parents @4 and @10.

Low specificity towards specific BlCa cell lines: We also selected two Nbs (**BlCaNb37** and **BlCaNb38**) that show no specificity towards one of the BlCa cell lines, meaning that the BL for all BlCa cell lines must be higher than zero during both SR1 and SR2.

Selection criterium 2: The total BLs of the Nb must be higher than 100, the AVR BL against BlCa cell lines must be 1.4-fold higher during SR2 compared to SR1. Also, we used the proprietary database to find the ratio between the AVR BL against BlCa tissue and bladder normal (BlN) tissue for the Nbs present in our database. Applying that the BL against BlCa tissue must be 1.5-fold higher than the BL against BlN tissue, we were able to retrieve 10 Nbs.

Selection criterium 2.1: From these 10 Nbs, we selected two Nbs (**BlCaNb42** and **BlCaNb43**) for which the BL against BlCa cell lines is much higher than non-BlCa cell lines, for which the BL increases during SR2 compared to SR1 and that originate from different family parents.

Additional selection: Applying a manual selection, we also selected four Nbs (**BlCaNb39**, **BlCaNb40**, **BlCaNb41** and **BlCaNb44**) that did not meet the above-mentioned criteria, but do show a BlCa-specific profile among the cell lines within the database, which can also be observed by the BL pattern in Figure 18.

Non-BlCa binding Nbs: We also selected Nbs that are not BlCa specific and can serve as negative controls during validation of our Nbs. For selection, we used the following criteria: the total BL must be higher than 200, the AVR BL to BlCa cell lines must be at least 2-fold lower than the AVR BL to non-BlCa cell lines during SR1 and 4-fold lower during SR2, the AVR BL against BlCa cell lines must be 2-fold lower during SR2 compared to SR1, while the AVR BL against non-BlCa cell lines must be 3-fold higher during SR2 compared to SR1. Applying these criteria, we narrowed the 300,000 Nbs down to 20 Nbs of which we selected the two Nbs (**nonBlCaNb1** and **nonBlCaNb2**) with the highest total BLs.

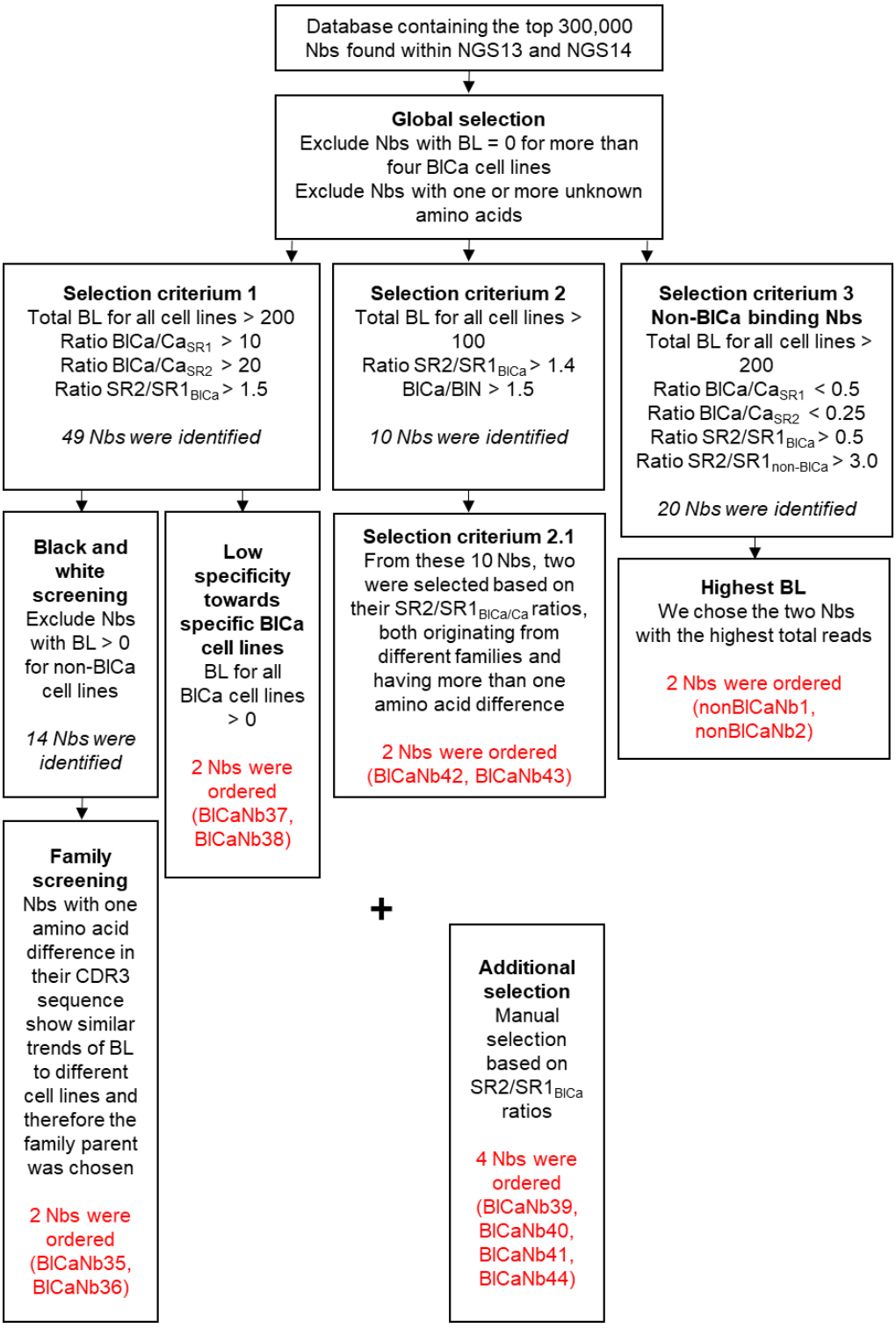


Figure 17 | Flow scheme of BlCaNb selection *in silico*. Using this flow scheme, we selected 12 Nbs from 300,000 Nbs in our NGS database. **SR2/SR1$_{BlCa}$** = Ratio between the AVR BL of Nbs to BlCa cell lines (T24, JON, J82, TCCSUP, RT112, 5637) during selections one and two. **BlCa/Ca$_{SR1}$** = Ratio between the AVR BL of Nbs to BlCa cell lines and non-BlCa cell lines during the first selection. **BlCa/Ca$_{SR2}$** = Ratio between the AVR BL of Nbs to BlCa cell lines and non-BlCa cell lines during the second selection. **BlCa/BlN** = Ratio between the AVR BL of Nbs to BlCa tissue and Bladder Normal tissue samples (This data was obtained from an Nb sequence library database generated in previous projects). **SR2/SR1$_{nonBlCa}$** = Ratio between the AVR BL of Nbs to non-BlCa cell lines (LNCaP, VCaP, MDAPCa2b, PC346C, B16, B16-PSMA, MCF7 and T47D) during selections one and two. **SR2/SR1$_{BlCa/Ca}$** = Ratio between BlCa/Ca$_{SR2}$ and BlCa/Ca$_{SR1}$. **BL** = binding level.

	Family parent	Total BL	5637 (1)	J82 (1)	JON (1)	RT112 (1)	T24 (1)	TCCSUP (1)	5637 (2)	J82 (2)	JON (2)	RT112 (2)	T24 (2)	TCCSUP (2)
BlCaNb35	4	680341	116	113160	107418	14042	7	100174	7	139560	173184	6012	7	184944
BlCaNb36	10	342273	70	62094	43150	2277	3	51984	8	110020	78421	1406	0	110633
BlCaNb37	182	11623	2545	11	171	1531	429	75	7424	4	77	3038	326	28
BlCaNb38	673	278	28	29	15	12	42	18	32	53	22	25	56	33
BlCaNb39	969	1087	67	30	72	163	89	67	140	27	93	383	99	73
BlCaNb40	19	216732	14287	5588	6840	7782	11179	6320	43743	7800	14994	18650	26055	11360
BlCaNb41	673	1744	140	38	56	75	109	60	296	119	195	225	264	100
BlCaNb42	13	857	46	11	40	18	18	27	160	77	249	159	49	31
BlCaNb43	6	610953	69880	86129	8099	1990	80285	61873	38671	96410	3341	654	74011	91727
BlCaNb44	9	1089	128	54	35	51	96	50	243	49	46	123	114	64
nonBlCaNb1	68	60393	220	38	145	96	38	42	59	21	102	59	23	19
nonBlCaNb2	107	28627	38	29	16	8	6	74	8	8	6	0	4	43

LNCaP (1)	PC346C (1)	MDAPCa2b (1)	VCaP (1)	LNCaP (2)	PC346C (2)	MDAPCa2b (2)	VCaP (2)	B16 (1)	B16-PSMA (1)	MCF7 (1)	T47D (1)	B16 (2)	B16-PSMA (2)	MCF7 (2)	T47D (2)
1782	916	66087	21	865	91	40144	2	715	244	118	37	11	7	31	4
283	189	15669	16	169	13	7643	2	326	47	51	7	8	9	4	4
0	7	2	0	0	0	1	0	6	3	0	46	0	0	0	4
0	0	1	0	0	0	0	0	6	8	3	0	0	0	1	0
2	24	10	1	0	4	1	0	0	0	5	53	0	0	4	47
1628	1422	6359	4342	1235	790	11623	5760	13973	26951	7657	6288	10527	13592	17135	11706
2	9	44	24	3	0	48	15	118	194	39	44	82	25	87	77
30	18	23	8	32	34	31	11	0	3	35	19	0	4	13	34
42603	8222	26624	8807	72932	1209	46103	4957	453	107	11958	43	3	315	4188	19
6	7	21	25	7	0	20	3	105	153	18	16	42	29	38	28
455	614	283	1028	429	626	174	1943	5974	14298	151	355	14361	64384	62	335
183	682	106	102	109	1365	174	150	3792	1881	207	127	29199	2568	179	140

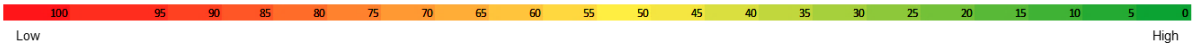


Figure 18 | The BLs of the 12 selected BlCaNbs against all Ca cell lines during SR1 (1) and SR2 (2). For each Nb, the family parent, total BL and BL against each cell line is depicted. Conditional formatting is used to generate a red-yellow-green colour scale with 50%-percentile as the midpoint.

From the selected BlCaNbs, we identified all Nbs with the selected CDR3 amino acid sequence within our NGS database. One CDR3 sequence could have been sequenced multiple times during NGS having slightly different CDR1 and/or CDR2 sequences. We retrieved the most commonly found Nb in the compilation of CDR3-identicals and checked the entire Nb sequence of our BlCaNbs for absence of stopcodons and frameshifts. The cDNA sequence is synthesized and cloned into pHEN phagemids vectors (Twist Bioscience). These phagemids contain an origin of replication (ori), a lac domain which is needed for gene regulation, the BlCaNb sequence with a Myc-6xHis-tag in frame with the pIII open reading frame, an f1 ori that is needed for the replication and packaging of the ssDNA of the phagemid [58], an ampicillin resistance domain and multiple digestion sites for restriction enzymes. The digestion sites within the BlCaNb sequence domain differ for each Nb, while the other restriction sites are present within all BlCaNbs, such as SfiI, NotI and NdeI. An example pHEN phagemid vector is observed in Figure 19.

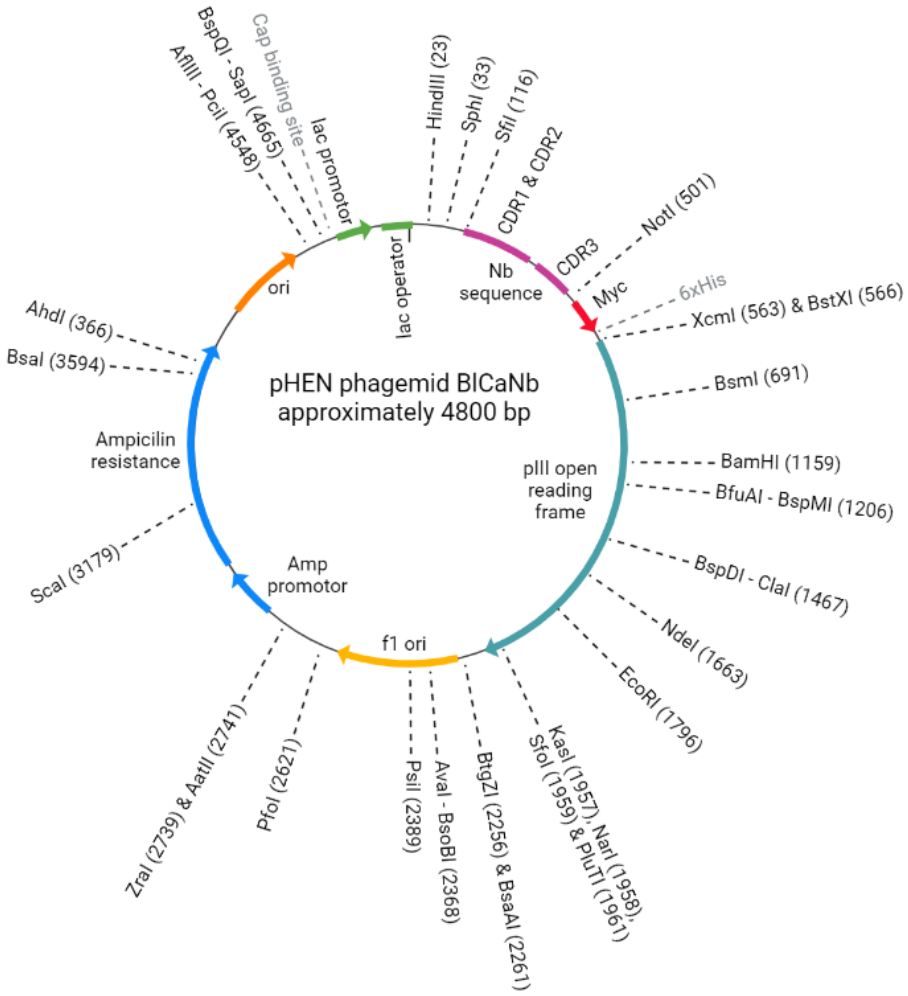


Figure 19 | pHEN phagemid BlCaNb. Example pHEN phagemid vector of our ordered BlCaNbs of approximately 4,874 bp (depends on the CDR length). The vector contains the Nb sequence, myc-tag, pIII open reading frame, f1 ori, Amp promotor, Amp resistance domain, ori and a lac domain consisting of a promotor and operator. Possible restriction sites are depicted as ----- Enzyme (location in bp). *Created with BioRender.com.*

4.6 Validation of BlCaNb pHEN phagemid vectors using restriction enzymes NotI and EcoRI

To verify the correct construction of the pHEN phagemid including our BlCaNb sequence, we performed a restriction enzyme assay where we cut our phagemid DNA using restriction enzymes NotI and EcoRI. These enzymes will cut the phagemid into two strands of respectively 3,500 and 1,300 bp lengths. When no Nb sequence is present, this restriction will lead to strands of 3,000 and 1,300 bps. Within this assay, we cut our 12 ordered BlCaNbs with both restriction enzymes and also added three extra samples, namely BlCaNb35 without restriction enzymes, BlCaNb37 when only cut with EcoRI and BlCaNb38 when only cut with NotI. After putting our samples on gel, we observed two bands for BlCaNb35 without restriction enzymes indicating the different supercoiled conformations of the phagemid. For BlCaNb37 and BlCaNb38, only cut with one restriction enzyme, we observed a clear band at height 5,000, indicating the single-stranded phagemid DNA. We observed three bands for

BlCaNbs 35, 36, 38-41, 43, 44, nonBlCaNb1 and nonBlCaNb2. The top band is around 5,000 bp, which indicates incomplete restriction of the phagemid DNA. The other two bands are at heights 3,500 and 1,300 bp indicating the two strands after complete restriction. For BlCaNb37 and BlCaNb42, we observed only two bands at heights 3,500 and 1,300 bp, meaning complete restriction of the phagemids by both NotI and EcoRI (Figure 20). Together with the quality control phagemid sequencing performed at Twist Bioscience, we conclude that all pHEN phagemids are correctly constructed and will be used for further validation.

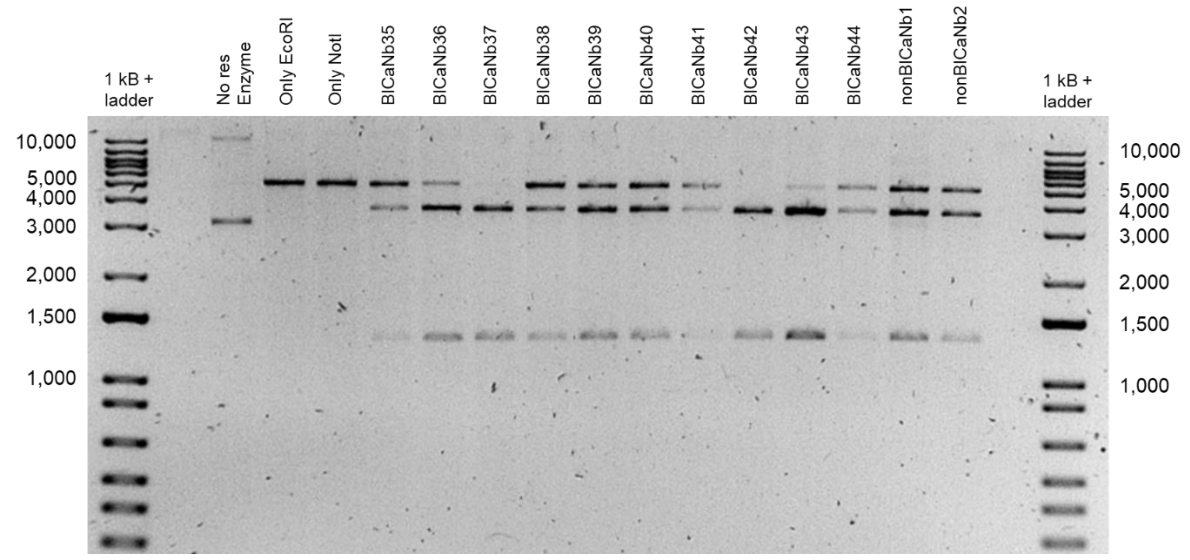


Figure 20 | Gel electrophoresis of BlCaNb pHEN phagemids after restriction enzyme digestion. 1 kB+ ladder = digested DNA including fragments ranging from 100 to 10,000 bps. The bp lengths are indicated on the right and left of the image. **No res Enzyme** = BlCaNb35 without restriction enzymes showing the supercoiled phagemid conformations. **Only EcoRI** = BlCaNb37 cut with restriction enzyme EcoRI. **Only NotI** = BlCaNb38 cut with restriction enzyme NotI. **BlCaNb35-44**, **nonBlCaNb1-2** = pHEN phagemids after restriction with NotI and EcoRI. This gel electrophoresis (1% agarose) was performed at 100 V for 120 min with 10 µL DNA sample (at different concentrations) and 2 µL Orange G loading buffer per well.

4.7 Validation of BlCaNbs

4.7.1 BlCaNb immunocytochemical staining on FroCellCAs

As a first validation step of our BlCaNbs, we performed immunocytochemical staining on FroCellCAs (see ‘Materials and Methods’ sections 3.9 and 3.10). In total, 21 Ca cell lines were air-dried onto glass slides and were incubated with phages that contained our Nbs fused to their pIII coat protein. As a control of this validation method, we incubated FroCellCAs with H6, NoBi and A7. After staining with H6, we observed positive staining for all Ca cell lines on the slide (Supplementary Data, section 6: Figure 40), while staining with NoBi led to zero staining for all Ca cell lines (Supplementary Data, section 6: Figure 42). For A7, we did not observe PSMA-positive staining which might be caused by the durability and stability of these phages (Supplementary Data, section 6: Figure 41). Previous stainings with A7 did show the expected binding. However as both the negative and positive control were successful, we continued with the analysis of the FroCellCAs. We analysed the slides using the

FroDA score to distinguish between cell line affinity for each Nb, the FroDA was not used to compare between different Nbs. The scores are normalized with the FroDA score for each cell line when stained with NoBi.

- For BlCaNb35, the highest four FroDA scores were for BlCa cell lines T24, TCCSUP, J82, RT112, and showed strong binding to the cell membrane. We also observed cell membrane staining for six out of 14 non-BlCa cell lines, while cell lines PC3 and DU145 were completely negative (Supplementary Data, section 6: Figure 43).
- For BlCaNb36, the top three FroDA scores were also BlCa cell lines (T24, J82, RT112), but also staining for five out of 14 non-BlCa cell lines was observed (Supplementary Data, section 6: Figure 44).
- For BlCaNb40 and BlCaNb41, the overall level of staining was low compared to BlCaNb35 and 36. Nevertheless, a clear DAB signal was visualized at the cell membrane of BlCa cell lines RT112, J82, T24 and TCCSUP, while for non-BlCa cell lines, we observed almost zero staining except for PC3 (Supplementary Data, section 6, Figures 49 and 50).
- For BlCaNb37, BlCaNb39 and BlCaNb42, only one out of six BlCa cell lines showed staining and also staining of PCa cell lines was observed (Supplementary Data, section 6, Figures 45, 47, 50). Therefore, these Nbs are not considered BlCa-specific.
- For BlCaNb38, we observed staining for three out of six BlCa cell lines, however, this staining was not localized at the cell membrane (Supplementary Data, section 6, Figure 46). We hypothesize that this staining was caused by experimental artefacts as also non-BlCa cell lines showed this pattern.
- Surprisingly, BlCaNb43 and BlCaNb44 showed zero staining against all cell lines (Supplementary Data, section 6, Figures 51 and 52).
- Lastly, we also incubated FroCellCAs with nonBlCaNb1, leading to staining of cell lines 22Rv1, J82, PC346C and TCCSUP, while staining with nonBlCaNb2 led to staining of cell lines 22Rv1, VCaP, PC346C, TCCSUP, 5637, MCF7 and H460 (Supplementary Data, section 6, Figure 53 and 54). Both nonBlCaNbs show affinity to both BlCa and other Ca cell lines and are therefore considered unspecific.

Overall, we conclude that BlCaNb35, 36, 40 and 41 show the best prospects concerning BlCa-specificity. They have high FroDA scores for most of the BlCa cell lines (Figure 21), show localized cell membrane staining and show low specificity towards non-BlCa cell lines.

Within our NGS database, we observed similar binding patterns of BlCaNb35 and BlCaNb36, related to their similarity in CDR3 sequences (1 of 17 amino acids difference). Their BLs within the NGS database have Spearman's correlation coefficient ($R_{spearman}$ = 0.9667, ρ (2-tailed) = 0.0000). When comparing the order of cell line affinity based on FroDA scores using Spearman's correlation, we also

observed a high correlation between these two Nbs ($R_{spearman}$ = 0.56509, ρ (2-tailed) = 0.00942). Additionally, we compared all Nbs based on their FroDA score against each cell line using Spearman's correlation and the results can be observed in Figure 22.

BlCaNb35									
T24	TCCSUP	J82	RT112	MCF7	HCC1806	5637	PC346C	JON	22RV1
36.9	27.5	26.5	21.0	20.9	19.3	18.6	18.3	17.6	16.5
HEP3B	COLO205	MDAPCa2b	HEK293	HUVEC	PC3	DU145	HU460	VCaP	LNCaP
15.0	13.3	12.4	9.6	0.0	-0.6	-2.4	-4.6	-7.5	-8.6
BlCaNb36									
T24	J82	RT112	PC346C	TCCSUP	22RV1	JON	HEP3B	VCaP	COLO205
30.3	20.3	18.8	18.2	17.4	14.9	11.8	11.0	9.6	8.6
HCC1806	HEK293	HUVEC	5637	HU460	DU145	LNCaP	MDAPCA2b	MCF7	PC3
6.4	6.4	3.8	3.2	2.8	-0.6	-0.6	-0.8	-1.0	-4.5
BlCaNb40									
J82	TCCSUP	T24	JON	COLO205	HEK293	HEP3B	HCC1806	MCF7	HUVEC
26.5	18.7	16.3	16.0	12.4	12.2	11.9	11.7	11.7	11.5
22Rv1	PC3	5367	PC346C	RT112	VCaP	MDAPCA2b	LNCaP	DU145	HU460
7.5	4.7	3.7	3.4	3.4	3.1	2.6	2.0	1.2	1.0
BlCaNb41									
J82	T24	22Rv1	MCF7	PC346C	TCCSUP	HCC1806	MDAPCA2b	VCaP	JON
22.0	8.7	7.7	7.2	7.2	6.9	5.3	4.6	4.3	3.6
COLO205	RT112	H460	Hep3B	DU145	PC3	LNCaP	5637	HEK293	HUVEC
3.3	3.1	2.4	2.3	1.4	1.2	0.5	0.2	-0.5	-4.1

Figure 21 | FroDA scores for BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 against 20 different cell lines. The scores were normalized by subtracting the FroDA score per cell line when stained with NoBi. The FroDA scores are shown from high to low (left to right).

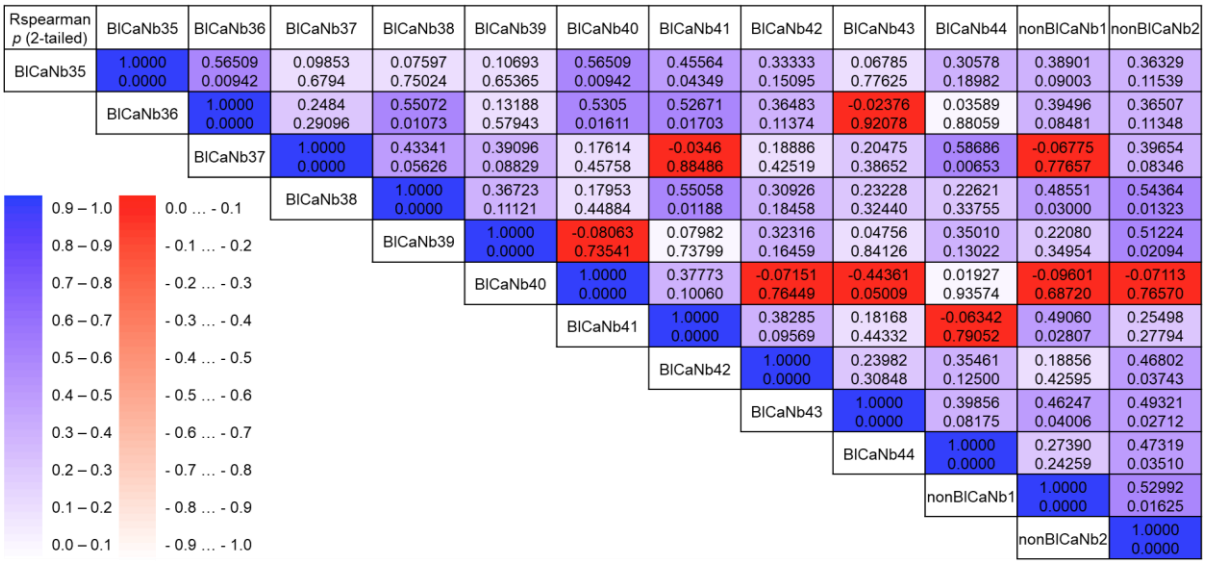


Figure 22 | Spearman correlation coefficients calculated between all BlCaNbs comparing FroDA scores. The upper value in each cell is $R_{spearman}$ and the lower value is ρ (2-tailed), where < 0.05 is considered statistically significant. A blue-red colour scale is used to indicate high and low correlations.

Furthermore, we compared the FroDA scores with Nb BLs during both SR1 and SR2. In total, 11 cell lines were present both on the FroCellCAs and within the NGS database. When focusing on SR2, we observed a low correlation for BlCaNb35 ($R_{spearman}$ = 0.18679, ρ (2-tailed) = 0.58236), for BlCaNb36 ($R_{spearman}$ = 0.10909, ρ (2-tailed) = 0.74951) and BlCaNb41 ($R_{spearman}$ = -0.16364, ρ (2-tailed) = 0.63069),

while BlCaNb40 showed a slight correlation ($R_{spearman}$ = 0.3098, ρ (2-tailed) = 0.35387). All Spearman correlation coefficients between FroDA scores and SR1/SR2 BL within the NGS database can be observed in Figure 23.

	Rspearman SR1 *p* (2-tailed)	Rspearman SR2 *p* (2-tailed)
BlCaNb35	0.04546 0.89443	0.18679 0.58236
BlCaNb36	0.13636 0.68931	0.10909 0.74951
BlCaNb37	-0.08736 0.79841	-0.22849 0.49919
BlCaNb38	-0.21610 0.52335	-0.03264 0.92410
BlCaNb39	0.29748 0.37431	0.34174 0.30366
BlCaNb40	0.30980 0.35387	0.30980 0.35387
BlCaNb41	-0.17273 0.61154	-0.16364 0.63069
BlCaNb42	-0.14253 0.6759	-0.16895 0.61947
BlCaNb43	-0.51378 0.10596	-0.47708 0.13785
BlCaNb44	0.02299 0.94651	0.14714 0.66595
nonBlCaNb1	-0.23235 0.49177	-0.25513 0.44897
nonBlCaNb2	0.0000 1.0000	-0.04100 0.90473

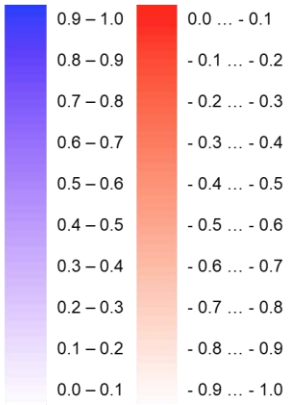


Figure 23 | Spearman correlation coefficients calculated for all BlCaNbs comparing the FroDA scores with the BLs within the NGS database. Within the figure, the $R_{spearman}$ and ρ (2-tailed) are depicted for each Nb comparing the FroDA score with BL for SR1 and SR2 within the NGS database. A blue-red colour scale is used to indicate high and low correlations. ρ-values < 0.05 are considered statistically significant.

4.7.2 BlCaNb40 might be a NECTIN4 binding Nb

In a previous project, performed by Jay Yee, HEK293 cells were transfected with lentiviral vectors leading to the expression of NECTIN4, UPK2 and UPK3B, which are BlCa marker proteins. First he performed a negative selection where the pre-selected libraries were panned against HEK293-WT cells, most of the Nb-phages binding to these cells were excluded from the library. Secondly, he performed three rounds of panning using the preBlC-libraries against HEK293 cells expressing NECTIN4, UPK2 and UPK3B. He found that UPK2 and UPK3B transfected HEK293 cells did not express the proteins on their membranes, but intracellularly. Therefore, to perform panning, he used beads coated with UPK2 and UPK3B lysates of the transfected HEK293 cells. After panning, the DNA was isolated and used to perform PCR followed by NGS. From this NGS data, another NGS database was generated containing the BLs of the top 300,000 Nbs. Within this database, we retrieved all 12 ordered BlCaNbs and the BLs can be observed in Figure 24.

We observed that the BLs against UPK2 and UPK3B after SR1 are high for all Nbs, while significantly decreasing after SR2 and even more after SR3. This similarity between Nbs and the steep decrease in BLs when performing multiple rounds leads us to the conclusion that these Nbs bind non-specifically to

the beads instead of the UPK2 and UPK3B protein coating. However, we observe high BLs for BlCaNb40 after all SRs, with ratio SR2/SR1 NECTIN4 = 1.95, suggesting that BlCaNb40 might be a NECTIN4 binding Nb. BlCaNb42, BlCaNb43 and BlCaNb44 also show high BLs during SR1, however, these significantly decrease during SR2 and SR3, and therefore they are not binding to NECTIN4.

BlCaNbs	Total frequency	NECTIN4 SR1 cells	NECTIN4 SR2 cells	NECTIN4 SR3 cells	UPK2 SR1 beads	UPK2 SR2 beads	UPK2 SR3 beads	UPK3B SR1 beads	UPK3B SR3 beads
BlCaNb35	282043	3332	250	0	115170	39917	1050	97689	191
BlCaNb36	151631	1500	163	11	57224	28214	1343	49545	440
BlCaNb37	2128	0	0	0	971	151	0	764	0
BlCaNb38	61	0	0	0	25	0	0	36	0
BlCaNb39	564	0	0	0	108	168	0	218	0
BlCaNb40	834530	37246	69006	40720	7914	3413	53	6657	19
BlCaNb41	1142	77	75	0	116	34	13	127	0
BlCaNb42	13156	2768	1626	1622	954	34	0	746	0
BlCaNb43	2547898	263629	248929	67832	101167	58631	3562	85121	708
BlCaNb44	3263	308	138	42	58	17	0	36	0
nonBlCaNb1	4864	295	0	0	1070	336	0	1037	0
nonBlCaNb2	18257	1500	275	42	2107	3279	1329	3310	900

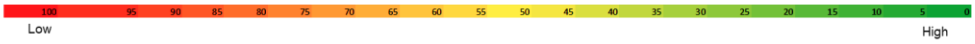


Figure 24 | BLs of our BlCaNbs within the NGS database generated after panning with preBlC-libraries for 1-3 rounds against HEK293 cells expressing NECTIN4 and beads coated with UPK2 and UPK3B. Conditional formatting is used to generate a red-yellow-green colour scale with 50%-percentile as the midpoint. BLs of Nbs after SR2 against beads coated with UPK3B were excluded due to PCR failure.

4.7.3 Higher Nb-phage concentrations lead to more unspecific binding of phages

After the first validation experiment using FroCellCAs, we observed specificity of BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 against BlCa cell lines and continued with further validation of these BlCaNbs. We aimed to increase the BlCa cell line staining by increasing the phage concentration from 5 x 10^{10} phages/mL to 10 x 10^{10} phages/mL (see 'Materials and Methods' section 3.10). For BlCaNb40 and BlCaNb41, we indeed observed an increased signal for the BlCa cell lines. For BlCaNb35 and BlCaNb36, the staining increased against all cell lines, which can be explained by more unspecific binding of phages when the concentration is increased (Supplementary Data, section 7: Figure 55 - 58). Therefore, for future experiments, we decided to keep the Nb-phage concentration at 5 x 10^{10} phages/mL.

4.7.4 Validation of BlCaNbs using whole cell and tissue enzyme-linked cell assay

With the four BlCaNbs, we performed enzyme-linked cell assays (ELCA) using A7, H6 and NoBi as control Nbs (see 'Materials and Methods' section 3.15). For the first assay, we included cell lines RT112, J82, 22Rv1, HCC1806, LNCaP and DU145, and used healthy bladder, lung, colon, salivary gland and stomach tissues. We measured the binding activity of the phages using an HRP-conjugated antiM13 antibody, by measuring the OPD-product absorption at 450 nm and normalized these per cell line/tissue by subtracting the OPD-product signal strength of empty wells. The cell/tissue ELCA plots can be found in Supplementary Data section 8.

We observed a very low binding activity of our BlCaNbs against tissues compared to cell lines. During the assay, we observed a rapid reduction of tissue pellets within the wells due to washing steps and this might be causing the lower binding of our Nbs to tissue. Due to this, we cannot draw conclusions on tissue binding and we need to repeat the assay performing solid-tissue ELCA (see 'Results' section 4.7.5).

When analysing the binding activity against the different cell lines (Figure 25), we observed high binding specificity for A7 against cell lines LNCaP and 22Rv1, which are PSMA-expressing cell lines. For NoBi, we observed low binding against all cell lines, whereas H6 showed binding towards all cell lines. All four BlCaNbs tested showed high binding activity against J82, and three out of four BlCaNbs showed high binding activity against LNCaP. We calculated the ratios between binding activity against BlCa cell lines and non-BlCa cell lines and found ratios > 1 for BlCaNb36 (BlCa/non-BlCa ratio = 1.536 (1.5 x 10^5 cells) & 1.451 (3.0 x 10^5 cells); ρ-value = 0.51570 & 0.87288). For BlCaNb40 and BlCaNb41, we observed high signals for both BlCa cell lines included, however also high signals were observed against the other cell lines. From this assay, we conclude that we observe specificity of our BlCaNbs towards J82, however, we repeated this assay using more cell lines (see 'Results' section 4.7.6 and 4.7.7).

	J82 1.5	RT112 1.5	DU145 1.5	LNCaP 1.5	22Rv1 1.5	HCC1806 1.5	BlCa/non-BlCa	ρ-value
NoBi	0.483	0.255	0.385	0.752	0.5015	0.4195	0.717	0.365081
A7	0.5005	0.4275	0.606	1.407	1.486	0.6035	0.452	0.594199
H6	1.3335	0.9785	1.472	1.9865	1.662	1.1715	0.735	0.208508
BlCaNb35	1.4955	0.9805	1.38	1.521	0.8625	1.147	1.008	0.970923
BlCaNb36	1.795	0.392	0.6845	0.638	0.5355	0.9905	1.536	0.448142
BlCaNb40	1.498	0.8985	1.2435	1.293	0.7285	0.905	1.144	0.600112
BlCaNb41	1.9445	1.0675	1.673	1.2555	0.6355	0.9105	1.346	0.418857
	J82 3.0	RT112 3.0	DU145 3.0	LNCaP 3.0	22Rv1 3.0	HCC1806 3.0		
NoBi	0.755	0.508	0.5125	0.968	0.4935	0.4525	1.041	0.905504
A7	0.561	0.441	0.6135	2.099	1.404	0.59	0.426	0.281127
H6	1.7905	1.374	1.327	2.402	1.678	1.7045	0.890	0.616632
BlCaNb35	1.744	1.294	1.559	1.875	0.97	1.8965	0.964	0.881051
BlCaNb36	2.6225	0.639	1.1185	1.0455	0.7165	1.6165	1.451	0.490708
BlCaNb40	2.001	1.33	1.489	1.898	0.9115	1.5935	1.131	0.631624
BlCaNb41	2.5	1.4175	1.4765	2.214	1.3455	1.4665	1.205	0.496098

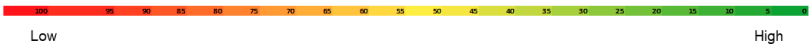


Figure 25 | Binding activity of our BlCaNbs, A7, NoBi and H6 against six Ca cell lines (J82, RT112, DU145, LNCaP, 22Rv1 and HCC1806). We performed ELCA with two different cell line concentrations: 1.5 = 1.5 x 10^5 cells and 3.0 = 3.0 x 10^5 cells. The values shown are the signal strengths of OPD at 450 nm normalized by subtracting the AVR signal strength of two empty wells (0.0905). Ratios were determined using the AVR signal strength for BlCa cell lines (J82 and RT112) divided by the AVR signal strength for non-BlCa cell lines (DU145, LNCaP, 22Rv1, HCC1806). Conditional formatting is used to generate a red-yellow-green colour scale with 50%-percentile as the midpoint. The ρ-values were calculated using the Student t-test for independent values (ρ-value (2-tailed) < 0.05 is considered significant).

4.7.5 Tissue-based enzyme-linked cell assay led to unquantifiable results

We analysed our BlCaNbs with ELCA using solid healthy tissues. We used similar tissues as before, namely colon, stomach, salivary gland and lung, and incubated our samples at three different concentrations (2.25x, 3.0x and 5.0x) with 5 x 10^5 phages/mL of NoBi, H6, A7, BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 (see 'Materials & Methods' section 3.15). We observed that NoBi is indeed negative (< 0.1 OPD signal strength) for colon, lung and stomach, whereas NoBi showed some signal for salivary gland. For H6, we observed relatively high signals for lung, stomach and salivary gland (> 0.4 for all tissue concentrations), whereas colon was lower (0.2 - 0.3). For A7, a PSMA-binding Nb-phage, we observed a very high signal for salivary gland (> 0.6 for all tissue concentrations). This is in line with previous research [59], where it has also been shown that the salivary glands show high absorbance of PSMA-targeting radionuclides-conjugated peptides. For BlCaNb35 and BlCaNb36, we observed high signals for colon and salivary gland, whereas lung and stomach are relatively lower. For BlCaNb40, we observed low signals for colon, lung and stomach, whereas the signal for salivary gland is again higher. Lastly, for BlCaNb41 we observed similar signals for all four tissues going from approximately 0.3 to 0.6 (Figure 26).

When performing tissue-based ELCA, we observed high inter-assay variation. We hypothesize that this is caused due to differences in protein content within each tissue sample and between different assays, and also due to the presence of blood and fat within the tissues. Dr. Raheleh Tooyserkani also tried to lyse the tissue and performed BCA protein quantification. Then coated a defined concentration of the total protein from each tissue onto 96-plate wells and repeated the ELCA assays. Unfortunately, again high inter-assay variation was observed, mainly according to non-specific binding of NoBi as well as A7 to normal tissue protein lysates. Also, we are not sure of the conformation of membrane proteins after tissue lysis, adding even more complexity. Up to now, we have not succeeded to estimate tissue protein concentration exactly, therefore drawing conclusions based on these tissue ELCAs is unreliable.

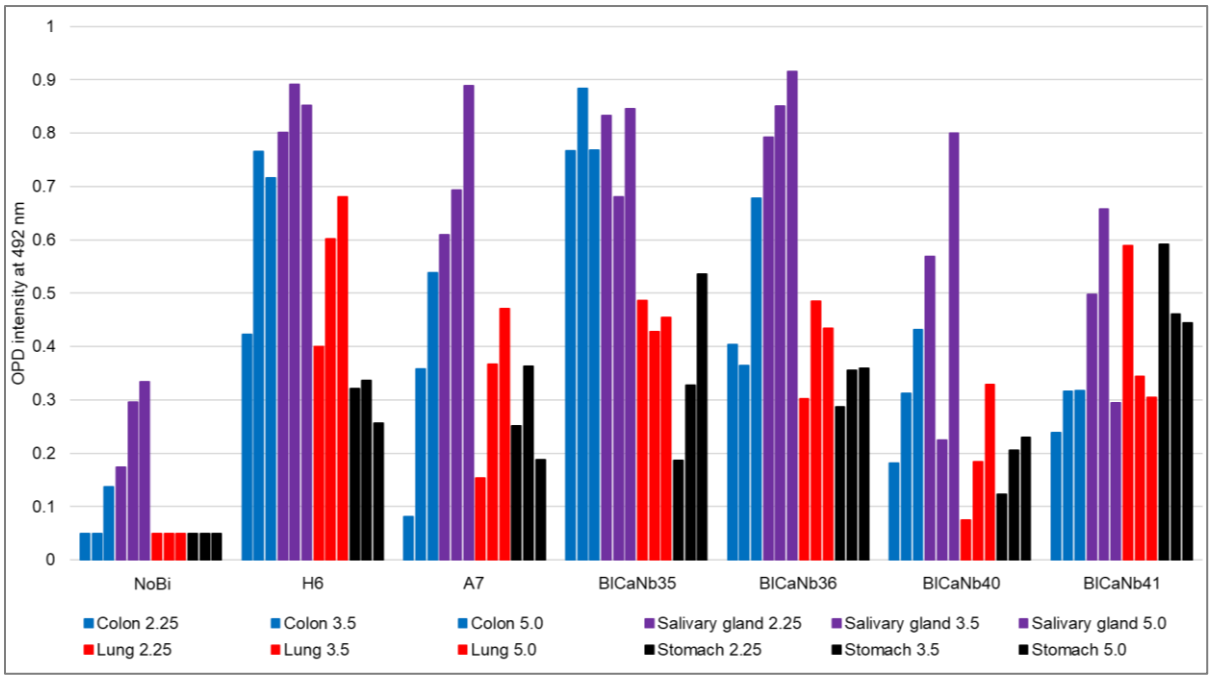


Figure 26 | Histogram of the binding activity of our BlCaNbs, A7, NoBi and H6 against four types of meshed tissue samples at three concentrations. We performed ELCA of our BlCaNbs against three different concentrations (low to high → left to right) of colon (blue), salivary gland (purple), lung (red) and stomach (black) tissues. On the y-axis, the signal strength of OPD at 450 nm is shown. On the x-axis, the different samples are shown.

4.7.6 BlCa-specificity validation with cell-line ELCA (1)

When performing ELCA with cell lines, we observed much more qualitative results and our control Nbs show similar binding throughout the different assays performed. We performed an ELCA with the following cell lines: five BlCa cell lines (RT112, T24, TCCSUP, J82, 5637), three PCa cell lines (DU145, 22Rv1, PC346C), two breast Ca cell lines (T47D, MCF7), a kidney Ca cell line (HEK293) and a colon Ca cell line (COLO205). Besides measuring the OPD signal strength for BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41, we again used H6, A7 and NoBi as control Nbs. For NoBi, we observed an OPD signal at 450 nm of $\mu = 0.35 \pm \sigma = 0.11$ (Figure 27 left), whereas, for H6, we observed $\mu = 2.35 \pm \sigma = 0.044$ (Figure 27 middle). For A7, we observed high signals for PSMA positive-cell lines 22Rv1 and PC346C (Figure 27 right). For BlCaNb35, we observed high signals for cell lines RT112, J82, TCCSUP and DU145, whereas all other cell lines show low signal strengths. For BlCaNb36, a similar pattern as for BlCaNb35 is observed. During the assay, the signal strength for cell lines RT112, J82 and T24 stained with BlCaNb40 could not be measured as they were out of the signal range of the plate reader. Therefore, we diluted those samples two-fold and measured them again. In Figure 28, one can observe the signal strengths of those diluted samples after correcting for dilution. Finally, for BlCaNb41, we observed high OPD signals for all BlCa cell lines, but also for five out of seven non-BlCa cell lines.

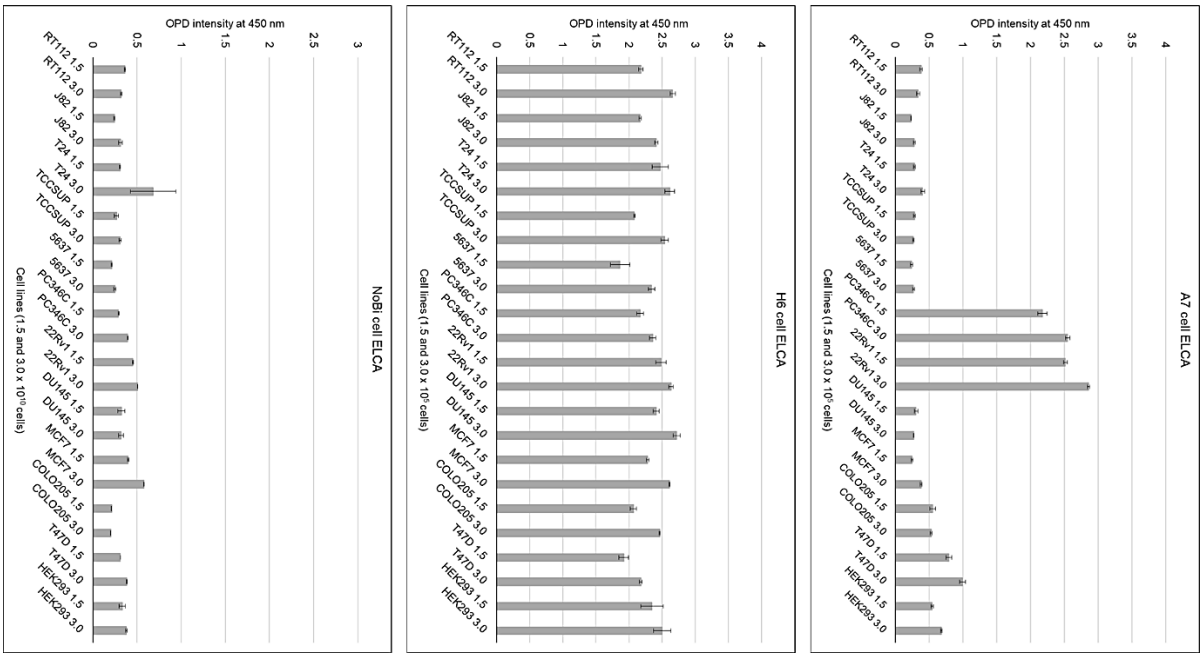


Figure 27 | NoBi, H6 and A7 Nb binding to different cell lines using enzyme-linked cell assay (ELCA) (1). Binding levels were tested after addition of 5 x 10^{10} phages/mL to 1.5 x 10^5 (1.5) and 3.0 x 10^5 (3.0) cells per cell line. All samples were included in duplicates, the averages are plotted and the error bar (standard deviation) is shown for each sample. On the x-axis, the signal strength of OPD at 450 nm is shown. On the y-axis, the different samples are shown.

This assay was based on a substrate reaction, where HRP converts OPD substrate into a yellow-orange product. The OD_{450nm} values of OPD-product are known to be reasonably linear between values 0.1-1.5, but linearity is lost with values above 2. Staying within this range can be a challenge, and adapting and monitoring the assay using a reaction-stop buffer at the right moment, sample dilutions and measuring OD_{450nm} at multiple time points can be implemented. The data depicted in Figure 28 should be interpreted with caution.

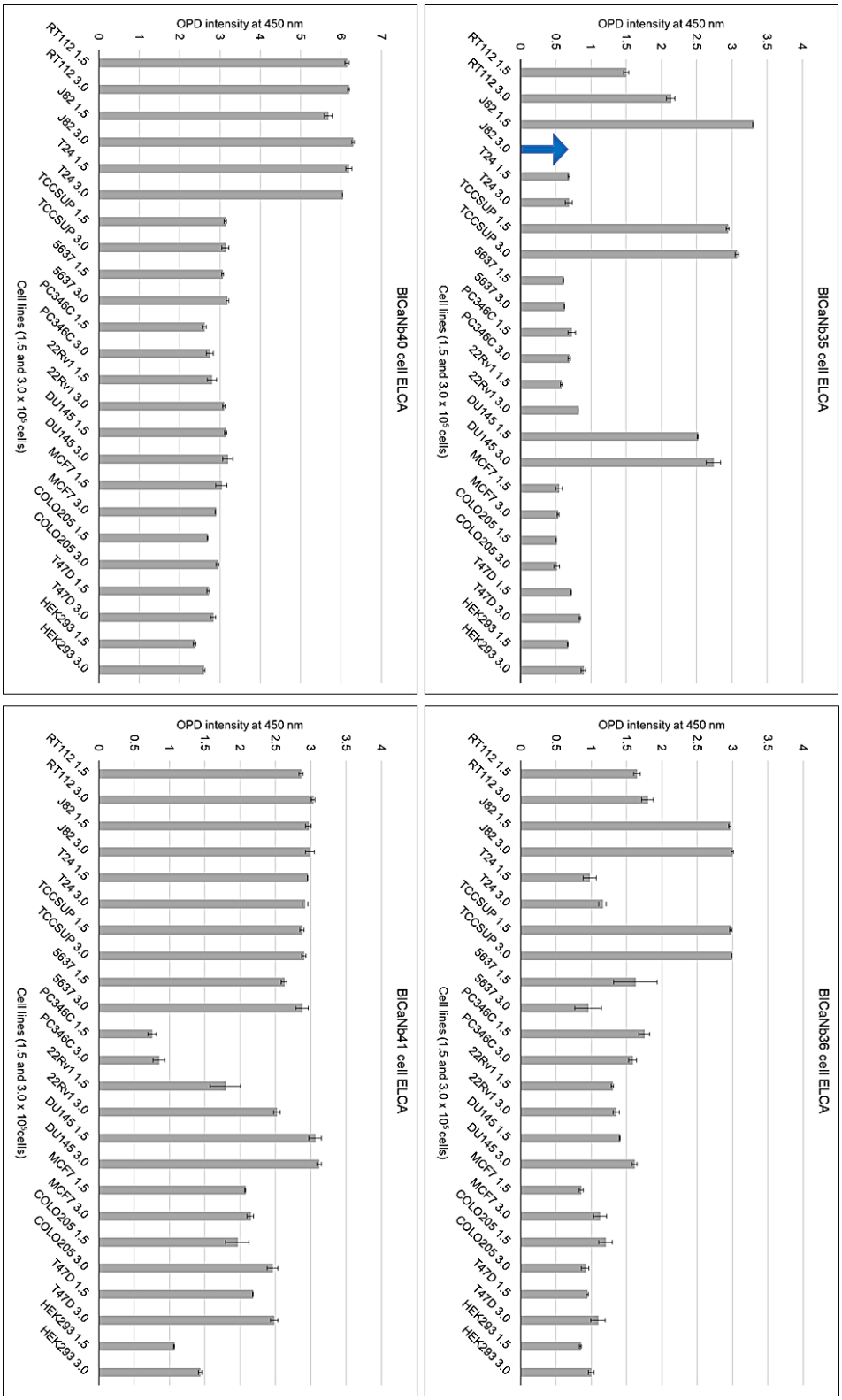


Figure 28 | BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 binding to different cell lines using enzyme-linked cell assay (ELCA) (1). Binding levels were tested after addition of 5 x 10^{10} phages/mL to 1.5 x 10^5 (1.5) and 3.0 x 10^5 (3.0) cells per cell line. All samples were included in duplicates during ELCA, the averages are plotted and the error bar (standard deviation) is shown for each sample. The blue arrow indicates a sample that could not be measured during ELCA and is therefore excluded from the plots. On the x-axis, the signal strength of OPD at 450 nm is shown. On the y-axis, the different samples are shown.

4.7.7 BlCa-specificity validation with cell-line ELCA (2)

As a final BlCaNb validation assay, we performed another ELCA in which we included six BlCa cell lines (RT112, T24, TCCSUP, J82, 5637, JON), 10 PCa cell lines (DU145, 22Rv1, PC346C, VCaP, LNCaP, BPH-1, PZ HPV-7, Ca HPV-10, RWPE-1, RWPE-2), a breast Ca cell lines (MCF7), a kidney Ca cell line (HEK293), two mouse cell lines (B16, B16-PSMA), an epithelial cell line (HUVEC) and a colon Ca cell line (COLO205). For this assay, we included less cells (3.0 x 10^4 and 6.0 x 10^4 cells) to make sure that we stay within the linear range of OPD measurements. Also, we measured at different time points and within the graphs we included the time points that were within the linear range. We used H6, A7 and NoBi as control Nbs and added 5 x 10^{10} phages/mL. For NoBi, we measured an OPD signal strength at 450 nm of $\mu = 0.0021 \pm \sigma = 0.017$ (Figure 31 left), whereas for H6, we measured an OPD signal strength of $\mu = 0.85 \pm \sigma = 0.28$ (Figure 31 middle). For A7, we again observed PSMA-specific high signals, where PSMA-positive cells (PC346C, 22Rv1, LNCaP, VCaP, B16-PSMA show high A7 signals (Figure 29 right). One may notice that HEK293 also shows a high OPD signal strength for A7, which can be explained by the fact that HEK293 are human embryonic kidney cells which have been shown to express PSMA [60-61].

We know that NoBi, H6 and A7 do not bind to B16 which is a skin melanoma mouse cell line. However, we do observe an OPD signal strength of approximately 0.4 for H6 and 0.3 for A7 against B16 cells. This leads to the conclusion that each phage has a different background signal caused by unspecific binding of the Nb-phage, as well as unspecific binding from phage coat proteins, and we may correct for this by subtracting the average signal strength of a specific Nb against B16.

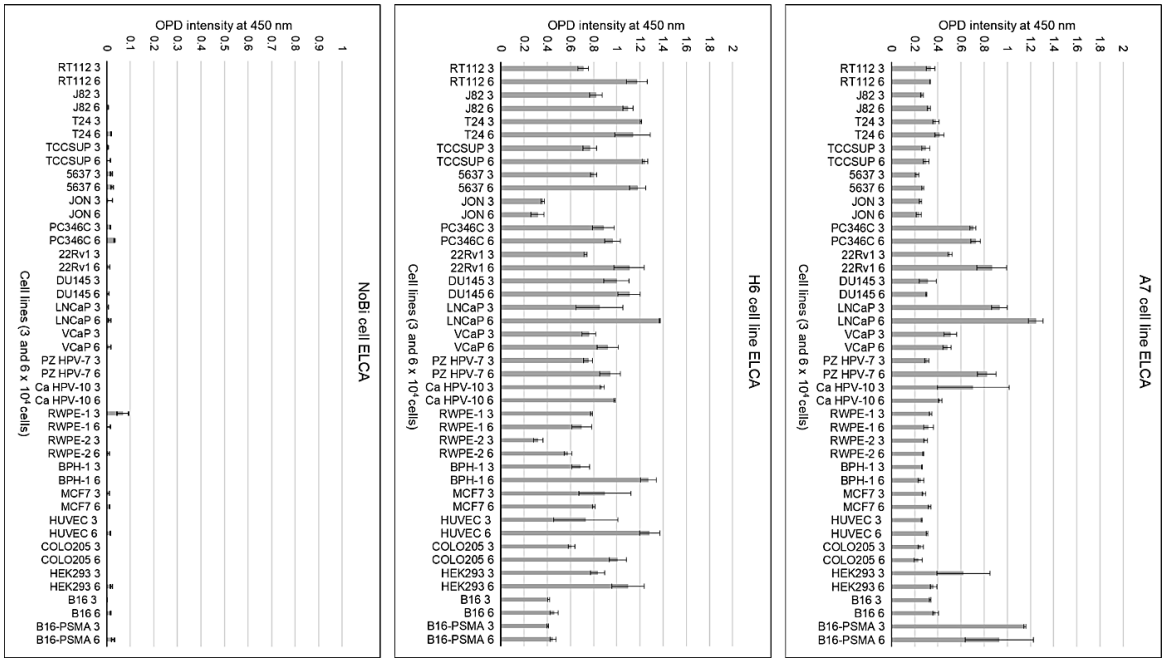


Figure 29 | NoBi, H6 and A7 Nb binding to different cell lines using enzyme-linked cell assay (ELCA) (2). Binding levels were tested after addition of 5 x 10^{10} phages/mL to 3.0 x 10^4 (3.0) and 6.0 x 10^4 (6.0) cells per cell line. All samples were included in duplicates, the averages are plotted and the error bar (standard deviation) is shown for each sample. On the x-axis, the signal strength of OPD at 450 nm is shown. On the y-axis, the different samples are shown.

We corrected the signal strength by subtracting the B16 signal strength and we calculated the ratios between ELCA signal strengths for BlCa and non-BlCa cell lines (Table 6). We found that BlCaNb36 and BlCaNb40 have BlCa/non-BlCa ratios for both 3.0 x 10^4 and 6.0 x 10^4 cells higher than 3.0, where the ratios for BlCaNb40 are also statistically significant. The ELCA plots without B16 correction can be observed in Figure 30 and with B16 correction in Supplementary Data Section 9.

Before correction for the averaged OPD signal strength against B16				
ELCA cell line assay	BlCa/non-BlCa (3.0 x 10^4 cells)	ρ-value	BlCa/non-BlCa (6.0 x 10^4 cells)	ρ-value
BlCaNb35	1.00	0.975855	1.23	0.148049
BlCaNb36	1.18	0.124377	1.25	0.172745
BlCaNb40	1.44	0.002075	1.84	0.000056
BlCaNb41	1.36	0.088487	1.34	0.075208
After correction for the averaged OPD signal strength against B16				
ELCA cell line assay	BlCa/non-BlCa (3.0 x 10^4 cells)	ρ-value	BlCa/non-BlCa (6.0 x 10^4 cells)	ρ-value
BlCaNb35	0.943	0.962067	2.51	0.180954
BlCaNb36	8.71	0.145169	3.17	0.195906
BlCaNb40	4.47	0.003151	3.86	0.000056
BlCaNb41	1.75	0.118184	1.55	0.101205

Table 6 | BlCa/non-BlCa ratios of our BlCaNbs when performing ELCA against Ca cell lines. BlCa/non-BlCa ratios were calculated between six BlCa cell lines (RT112, T24, TCCSUP, JON, 5637 and J82) and 16 non-BlCa cell lines (PC346C, 22Rv1, DU145, MCF7, COLO205, VCaP, LNCaP, HUVEC, HEK293, PZ HPV-7, Ca HPV-10, RWPE-1, RWPE-2, BPH-1, B16 and B16-PSMA). Both BlCa/non-BlCa ELCA signal strength ratios for 3.0 x 10^4 and 6.0 x 10^4 cells can be seen before and after correction for the B16 ELCA signal strength per Nb. The ρ-values were calculated using Student t-test for independent values (ρ-value (2-tailed) < 0.05 is considered significant).

Comparing ELCA signal strengths between BlCaNbs, we again observed a very strong relationship between BlCaNb35 and BlCaNb36 with $R_{spearman}$ = 0.73968 (Figure 30). Again, we performed a comparative analysis using the Spearman correlation coefficient between FroDA scores from a previous assay (see 'Results' 4.7.1) and ELCA signal strengths (Figure 31), and between BLs within the NGS database and ELCA signal strengths for BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 (Figure 31). Between FroDA scores and ELCA signal strengths, there was an overlap of 11 cell lines (RT112, J82, 5637, T24, TCCSUP, PC346C, 22Rv1, DU145, MCF7, COLO205 and HEK293), while between the ELCA and NGS database, we had an overlap of eight cell lines (RT112, J82, 5637, T24, TCCSUP, PC346C, T47D and MCF7). We compared the NGS database BLs after SR2 with the ELCA signal strengths for 3.0 x 10^4 and 6.0 x 10^4 cells, and observed the strongest relationship for BlCaNb35 with $R_{spearman}$ = 0.48136. Between the FroDA scores and the ELCA signal strengths, we observed the strongest relationship for BlCaNb40 with $R_{spearman}$ = 0.45104. Again, the overall relation between ELCA signal strengths and BLs within the NGS database is stronger compared to the FroDA scores versus NGS database BLs.

Rspearman *p* (2-tailed)	BlCaNb35	BlCaNb36	BlCaNb40	BlCaNb41
BlCaNb35	1.0000 0.0000	0.73968 0.0000	0.1826 0.23548	0.07907 0.60992
BlCaNb36		1.0000 0.0000	0.24455 0.10963	0.15335 0.32032
BlCaNb40			1.0000 0.0000	0.47444 0.00114
BlCaNb41				1.0000 0.0000

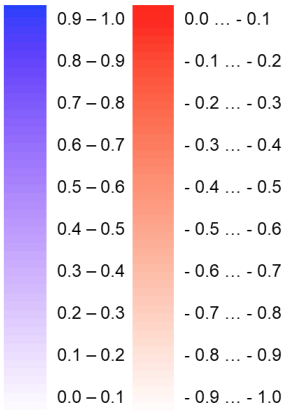
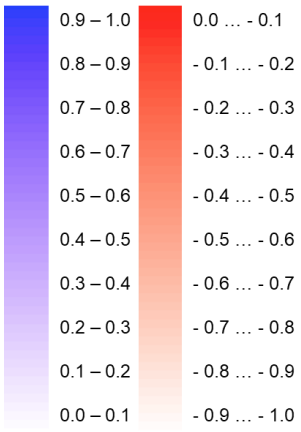


Figure 30 | Spearman correlation coefficients calculated comparing the ELCA signal strengths against different cell lines between BlCaNbs. The upper value in each cell is $R_{spearman}$ and the lower value is ρ (2-tailed), where < 0.05 is considered statistically significant. A blue-red colour scale is used to indicate the correlation strength.



	Nb database BL during SR2 vs ELCA	FroDA vs ELCA
	Rspearman *p* (2-tailed)	Rspearman *p* (2-tailed)
BlCaNb35	0.48136 0.01724	-0.1756 0.35332
BlCaNb36	0.27948 0.18597	0.04372 0.81858
BlCaNb40	0.25109 0.23661	0.45104 0.01236
BlCaNb41	0.17611 0.41041	-0.08914 0.63949

Figure 31 | Spearman correlation coefficients between the ELCA signal strengths and NGS database BLs during SR2, and between ELCA signal strengths and FroDA scores. Between the FroCellCa assay and the ELCA, there was an overlap of 16 cell lines (RT112, J82, 5637, T24, TCCSUP, JON, PC346C, LNCaP, VCaP, HUVEC, MCF7, 22Rv1, DU145, HEK293, COLO205 and B16), while between ELCA and the NGS database there was an overlap of 12 cell lines (RT112, J82, 5637, T24, TCCSUP, JON, PC346C, VCaP, LNCaP, B16, B16-PSMA and MCF7). The upper value in each cell is $R_{spearman}$ and the lower value is ρ (2-tailed), where < 0.05 is considered statistically significant. A blue-red colour scale is used to indicate the correlation strength.

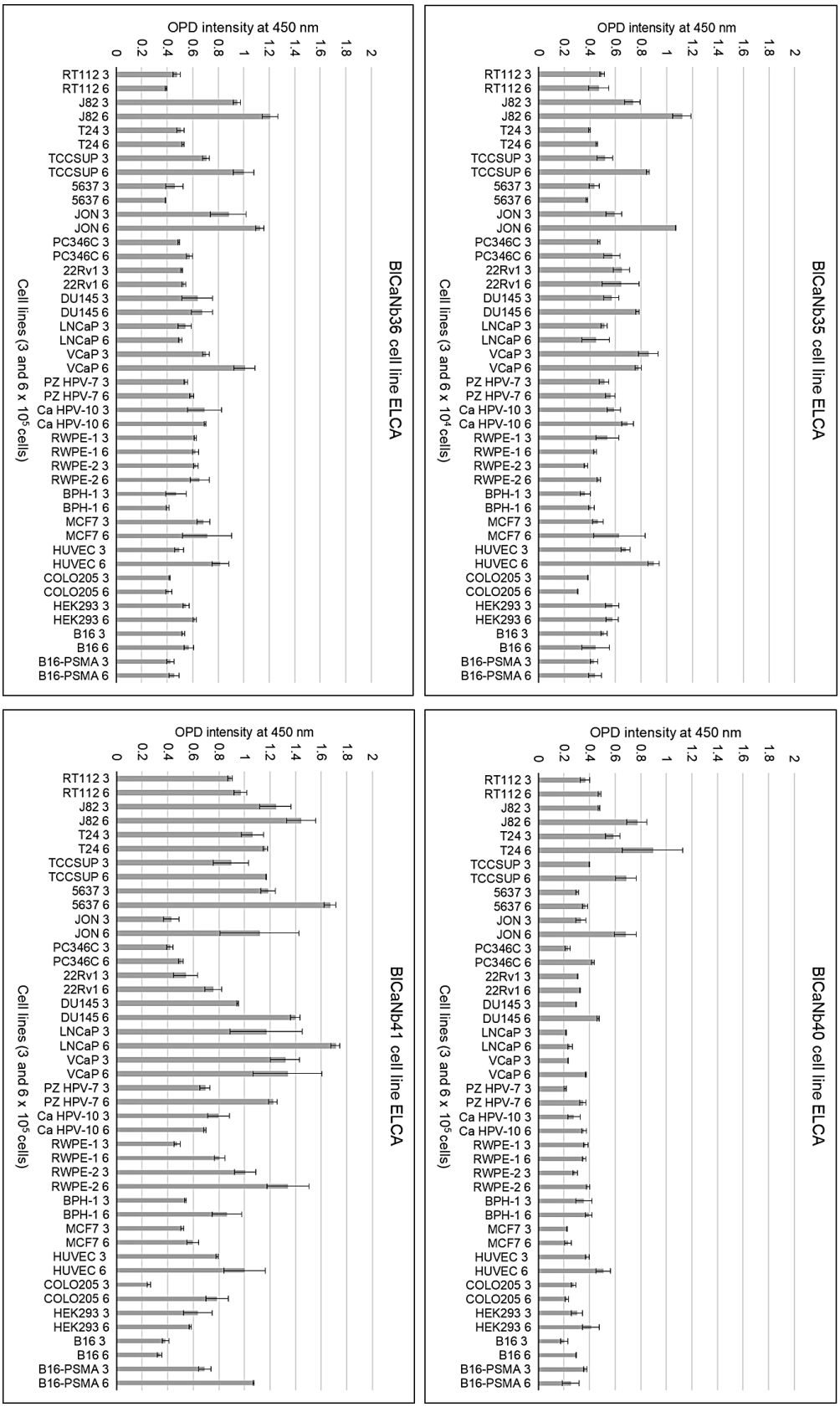


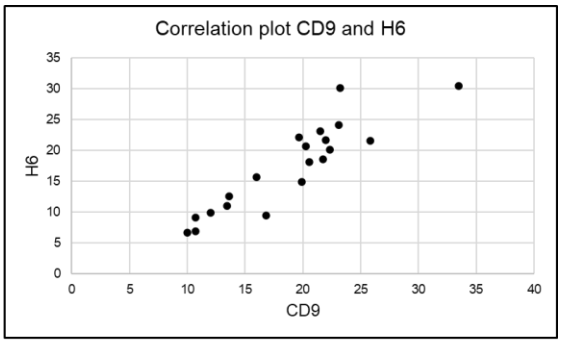
Figure 32 | BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 binding to different cell lines using enzyme-linked cell assay (ELCA) (2). Binding levels were tested after addition of 5 x 10^{10} phages/mL to 3.0 x 10^4 (3.0) and 6.0 x 10^4 (6.0) cells per cell line. All samples were included in duplicates during ELCA, the averages are plotted and the error bar (standard deviation) is shown for each sample. On the x-axis, the signal strength of OPD at 450 nm is shown. On the y-axis, the different samples are shown.

4.8 Validation of H6 binding to CD9 membrane proteins

During this project, we used H6 many times as a positive control Nb as it binds to CD9 cell membrane proteins. In the following assay, we compared the binding of anti-CD9 and anti-CD63 antibodies to H6 binding after performing immunocytochemical staining of FroCellCAs (see 'Materials and Methods' section 3.10). We calculated the FroDA scores against each cell line when stained with 5 x 10^{10} phages/mL of H6 and 2.5 ng/µL anti-CD9 / anti-CD63 antibodies (see 'Materials and Methods' section 3.16).

When comparing the FroDA scores per cell line for anti-CD9 antibody and H6, we observed a strong correlation between the two with a Spearman's correlation coefficient of $R_{spearman}$ = 0.77946 with ρ = 0.00005 (Figure 33). Between anti-CD63 antibodies and H6, we observed no correlation where $R_{spearman}$ = 0.33724 with ρ = 0.13739. In conclusion, we further corroborated that H6 binds CD9.

Cell line	CD9 FroDA score	H6 FroDA score
PC3	34	22
LNCaP	26	30
22RV1	23	18
Hep3B	23	7
DU145	22	19
Colo205	22	9
MDAPCa2b	22	20
PC346C	21	30
J82	21	11
VCaP	20	15
Hek293	20	9
T24	20	10
HUVEC	17	24
JON	16	22
HCC1806	14	22
MCF7	13	21
H460	12	7
5637	11	16
RT112	11	23
TCCSUP	10	13



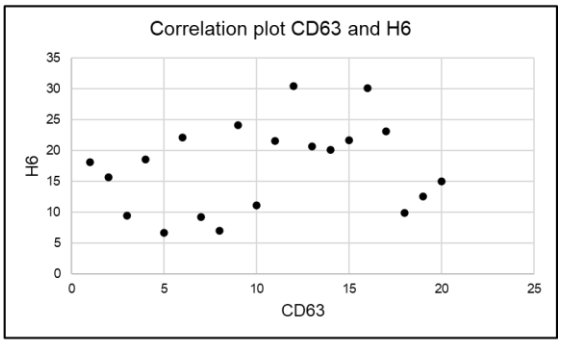


Figure 33 | H6 Nb and anti-CD9/anti-CD63 Ab binding comparison through FroCellCA staining. (Left) FroDA scores are shown for 20 different cell lines after staining with CD9-antigen and H6 Nbs. Conditional formatting is used to generate a red-yellow-green colour scale with 50%-percentile as the midpoint. (Top Right) Correlation plot between CD9 FroDA scores (x-axis) and H6 FroDA scores (y-axis) per cell line. (Bottom Right) Correlation plot between CD63 FroDA scores (x-axis) and H6 FroDA scores (y-axis) per cell line.

5 Discussion and Conclusion

We set out to gain more understanding about the effect of multiple selection rounds and a pre-selection of phage libraries when searching for BlCa-specific Nbs. Also, using an optimized panning procedure, we aimed for the identification of BlCa-specific Nbs.

5.1 Is pre-selection of Nb phage libraries better in identifying BlCaNbs?

During this project, we pre-selected three phage libraries fusing six BlCa cell lines. We tested this smaller sub-library of BlCa binding Nb-phages with immunocytochemical staining. With this, we showed that pre-selection increases the binding of these libraries to all Ca cell lines, not specifically BlCa cell lines. In the future, it would be important to also check Ca cell line affinity of these libraries with more quantitative assays, such as ELCA. We also found that BlCa-specific Nbs selected from the original NGS database have increased in BLs within the pre-selected library, leading to the conclusion that pre-selection is beneficial when searching for Ca-specific Nbs. Pre-selection can be performed against each cell / tumour type to improve identification of Ca-specific Nbs and we recommend this in future searches for Nbs specific against different Ca tumour types.

Our pre-selection is considered a positive selection and one could consider additionally performing a negative selection before selection against the Ca cell lines of interest. During this negative selection, the libraries can be panned against a mixture of healthy cell lines and most of the Nb-phages binding to these healthy cells are then excluded from the library. It is impossible to take out all non-specific binders, however, we believe that a combination of negative and positive selection could improve Nb selection even more. However, one must keep in mind that the more rounds of panning are performed, the higher the chance of losing potentially specific-Ca binding Nbs.

5.2 Is selection based on two panning rounds an improvement over a single selection round?

We performed two rounds of panning and showed that Ca-specific Nbs increase in BL during the second selection compared to a single selection. We have proven this as:

- Previously selected Nbs were validated through immunocytochemical staining and have shown not to be specific for BlCa, with $SR2/SR1_{BlCa} < 1$.
- Potential-PSMA binding Nbs have a $SR2/SR1_{B16\text{-}PSMA} > 1$, while non-PSMA binding Nbs decrease in BL against the B16-PSMA cell line during the second selection.
- Nbs with $SR2/SR1_{BlCa} > 2.0$, have higher BLs against the AVR BlCa cell lines compared to the AVR PCa cell lines during both SR1 and SR2.
- Nbs with $SR2/SR1_{BlCa} < 0.5$, have higher BLs against the AVR PCa cell lines compared to BlCa cell lines during both SR1 and SR2.

- Binding patterns of Nbs become less dispersed during SR2 and show a more cell line-specific pattern.

Overall, we have shown the benefit of performing a second selection when searching for BlCa-specific Nbs. In future pannings, it is highly recommended to perform at least two selection rounds as it eliminates many non-specific binding Nbs. In literature, three panning rounds for selecting Nbs against specific proteins is common practise. Whether a third round outweighs the sequencing and materials costs and extra time investment, can be investigated.

5.3 We identified novel BlCaNbs with a preferential binding to BlCa cell lines

We generated an NGS database containing the top 300,000 Nbs selected with an adapted version of the phage-display technique where we performed two consecutive panning rounds using pre-selected libraries against a selection of Ca cell lines. We selected 10 BlCa-specific Nbs and two non-specific Nbs. Validation of these Nbs was performed using immunocytochemical staining of FroCellCA slides and cell line/tissue ELCAs.

After FroCellCA analyses, we observed BlCa-specificity of BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41, whereas other BlCaNbs showed no specificity towards BlCa cell lines. We observed high similarity in binding patterns of BlCaNb35 and BlCaNb36 during both ELCA and immunocytochemical staining. This correlation was also observed within the NGS database and can be explained by their one amino acid difference in the CDR3 sequence. The amino acid changes are not expected to be disruptive as they interchange small amino acids Glycine to Alanine. This suggests that both Nbs bind to the same epitope and antigen. As similar binding was observed through all assays, it substantiates the robustness of our assays and the reproducibility of our results.

With ELCA, we performed experiments using both cell lines and tissue. Due to high inter-assay variation, caused by differences in protein content, and unspecific binding of Nb-phages, tissue ELCA was considered unreliable. Although BlCaNb40 showed the lowest binding to healthy tissue, it is difficult to conclude on tissue binding based on these ELCA assays. Next to tissue ELCA assays, we also performed whole cell line ELCA, where we measured the affinity of BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 to both BlCa and non-BlCa cell lines. We found that BlCaNb36 and BlCaNb40 showed the highest BlCa specificity. Again, we observed high correlation between BlCaNb35 and BlCaNb36 cell line binding. Using an NGS database generated by student Jay Yee, we also found that BlCaNb40 showed high BLs towards NECTIN4-expressing HEK293 cells, a BlCa marker protein. Whether BlCaNb40 indeed targets NECTIN4, needs to be validated using binding assays (ELCA, FroCellCA) with wild type HEK293 and NECTIN4-expressing HEK293 cells and subsequent epitope mapping. For BlCaNb41, we observed low specificity towards BlCa cell lines compared to the other BlCaNbs when validated using cell line ELCA.

Overall, based on the FroCellCA and ELCA assays, we have shown specific binding of BlCaNb35, BlCaNb36 and BlCaNb40 to BlCa cells. In the future, it would be important to continue validation of these Nbs and to exclude either BlCaNb35 or BlCaNb36 due to their binding similarity.

5.4 Assay development and comparisons

We compared FroDA scores and BLs within the NGS database for all BlCaNbs. We observed a low degree of correlation which can be explained by the fact that the molecular structure of surface molecules might be altered during the process of air-drying [62] when making the FroCellCA slides. On the contrary, when comparing ELCA signals with BLs of the NGS database, we observed high correlation for our BlCaNbs. For both the generation of the NGS database and ELCA assays, cell lines were processed in a hydrated state in which protein conformation is shown to be more stable than in shortage of water [62]. From this, we conclude that ELCA is more suitable for more accurate quantitative screening of Nb specificity, whereas FroCellCA can be used for quick screening.

During this project, we exploited and optimized the protocol of immunocytochemical staining of FroCellCA slides to improve the consistency of the staining strength of different Nbs. Nbs A7, H6 and NoBi proved this consistency, where NoBi always lead to zero staining, A7 to specific PSMA staining and H6 to CD9 staining. For very specific high affinity Nbs such as A7, differences in staining are prominently detected. This strong effect was unfortunately not observed for our BlCaNbs, which likely makes the FroCellCA less suitable for detection of small differences in Nb binding.

We used the FroCellCA assay as it has the advantage of being straightforward, cheap and quick, compared to for example fluorescent staining. We came up with a more quantitative method to analyse staining of FroCellCA slides, which we named the FroDA score. We found that FroDA scores correlated with Nb binding, where higher scores related to stronger binding. However, we still need to optimize this quantification method as it is not robust. We observed that slightly different microscope settings such as light, imaging thresholds and exposure time of the camera have a huge impact on the scores. We solved this problem by measuring all slides from one experiment in a single batch, thereby reducing experimental set-up differences. We also came to the conclusion that FroDA scores can be used to compare Ca cell line staining on one slide with one Nb, and poorly for the comparison between staining of different Nbs. Another problem that we encountered, was the difficulty to consistently load 5,000 cells per spot, leading to variability in cell density. Initially, more cells lead to more staining of Nbs. However, when cells become too tightly packed, less surface is available for the Nb-phages to bind to, leading to impeded phage binding. Therefore, it would be important to come up with a normalization factor that compensates for the difference in cell density or to adapt the protocol to improve cell density consistency.

For some BlCaNbs, we observed unspecific staining against all Ca cell lines. This might be caused by unspecific binding of M13 phage backbones to the cells or the glass slides. As we aim to screen and

validate many Nbs, we adapted all our assays to work with Nb-phages to prevent from expression and purification of each individual Nb; a labour-intensive procedure. In addition, using Nb-phages has the advantage that (i) five Nbs are present on one phage, increasing affinity as compared to monomeric purified Nbs and (ii) the huge phage has many (~2700) binding sites for the anti-M13 antibody directed against the main coat protein pVIII, which results in amplification of detection signals in our assays. To reduce unspecific binding caused by the M13 phage, one might have to produce the Nb protein without phages and validate specificity. Preferably, we do this only for a few highly promising candidates.

It is important to consider that Nb validation was not always performed on the same cell lines. We aimed to include similar cell lines for the different assays performed in this project. However, this was hampered due to cell line availability. This inconsistency hampered comparison of specific cell line affinity between the different assays.

5.5 Future perspectives

In total, we included six different BlCa cell lines during the entire project: RT112, J82, JON, T24, TCCSUP and 5637. It would be interesting to compare these cell lines on their RNA expression levels to see which genes are expressed and whether they are enhanced in BlCa, correlate this to BlCaNb BLs, and investigate how they play a role during BlCa development. We might also consider including more BlCa cell lines, such as KU1919 and 253J, to increase the genetic patterns against which we are screening our Nbs. Ultimately, we want to screen our libraries against the broadest possible pool of Ca antigens in which we include all Ca cell lines from the Cancer Cell Line Encyclopaedia (CCLE) [63]. And not only screen, but also generate an NGS database by performing panning against as much cells as possible.

For the top candidates, BlCaNb35, BlCaNb36 and BlCaNb40, we observed that none binds all BlCa cell lines. It is extremely unlikely one will find a Nb that specifically recognizes all BlCas in all patients: 100% sensitivity. One good Nb is a nice start, but a whole panel of different Nbs recognizing different BlCa subtypes are needed to be able to treat the majority of BlCas. Once BlCaNb35, BlCaNb36 and BlCaNb40 have shown to be correctly selected, the procedure of selecting Nbs from the NGS database, fill the database with more pannings against cell lines / tissue and even perform this using novel Nb phage display libraries, should be repeated to identify more BlCaNbs.

With our BlCaNb selection, we focused on Nbs that have high BLs against BlCa cell lines and low to other Ca cell lines. With this selection method, we exclude Nbs that are Ca-specific in general as we aim for BlCa-specificity. We might have eliminated the best Nbs that binds to many Ca cells and could improve Ca treatment in general. Which selection criteria do we use to select Nbs from our NGS database and what is the result of our selection are all questions one must take into account. Using the specificity of our BlCaNbs tested with FroCellCA and ELCA, we still need to reflect on the criteria on

which we selected our Nbs from the database in the first place such that we can improve selection from the NGS database in the future.

One of the most important questions left unanswered during this project is: Do our BlCaNbs bind to healthy tissue? In the end, we want our Nbs to bind as little as possible to healthy tissue to minimalize toxicity of our therapeutic Nbs. Previously, formalin-fixed paraffin-embedded (FFPE) tissues in the form of tissue microarrays (TMAs) were tested for validating Nb-phage binding. Unfortunately, it has been shown that this fixation method alters the protein structural organization and eliminates Nb binding, even upon different methods of antigen retrieval [64]. In this project, we utilized frozen cells and tissues to retain Nb-binding to the antigens. One complication is the fact that freeze-thawing disrupts membranes of cells, resulting in leakage of cytoplasm and loss of well-defined tissue histology. Checking for healthy tissue binding becomes cumbersome as generating frozen tissue sections and frozen TMAs is laborious and expensive. Currently, the laboratory is testing alternative fixation methods (e.g. TISPA) that retain 3D protein conformation and tissue integrity.

Using our pipeline, we do not know what the specific target is of our Nbs. To explain the Nb's specificity, knowing the antigen is essential. Various methods are available to map the epitope and the antigen, including mass spectrometry of bound protein, protein arrays and transmembrane protein cDNA expression screening. If the Nb recognizes a complex protein structure or glycosylation, finding the epitope might become a real challenge. Based on the results, experiments should be initiated to identify the targets of BlCaNb36 and BlCaNb40.

In the future, the Nbs should be used for targeted delivery of cytotoxic drugs or radionuclides that kill Ca cells or for Ca detection. The cell-targeted delivery of these drugs reduces toxicity to healthy neighbouring cells and increases the effectiveness. The most straight forward method is to conjugate the Nb to cytotoxic drugs (ADCs) such as monomethyl auristatin E [65] or to a radionuclide such as ^{177}Lu. In addition, Nbs can be incorporated within NPs that selectively deliver chemotherapy to BlCa cells, such as gemcitabine and cisplatin. Ca detection can be achieved through the conjugations of Nbs with radionuclides for PET/SPECT imaging or fluorescent dyes for tumour detection during surgery.

All in all, we have shown the potential of our adapted biopanning method using pre-selected libraries and two consecutive rounds of panning in the search for BlCa-specific Nbs. Using this method, we have found three potential BlCa-binding Nbs: BlCaNb35, BlCaNb36 and BlCaNb40. Future-wise, our Nbs need to be validated using tissues, tumour organoids, *in vivo* models and finally in human patient studies. We believe that our Nbs have the potential to become key players in the detection and treatment of BlCa.

Bibliography

[1] O. Ifeany, "A review on Bladder Tumor Antigens," *Cancer Therapy & Oncology International Journal,* pp. 1-12, 12 February 2018.

[2] "Johns Hopkins Medicine," n.d.. [Online]. Available: https://www.hopkinsmedicine.org/health/wellness-and-prevention/anatomy-of-the-urinary-system/.

[3] Y. Chan, S. Sandlin, E. Kurzrock and S. Osborn, "The current use of stem cells in bladder tissue regeneration and bioengineering," *Biomedicines,* 2017.

[4] K. Andersson and K. McCloskey, "Lamina propria: the functional center of the bladder?," *Neurourology and urodynamics,* pp. 9-16, January 2014.

[5] P. Khandelwal, S. Abraham and G. Apodaca, "Cell biology and physiology of the uroepithelium," *American Journal of Physiology-Renal Physiology collections,* pp. 1447-1501, Dec 2009.

[6] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell,* vol. 144, no. 5, pp. 646-674, 4 March 2011.

[7] D. Hanahan and R. Weinberg, "Hallmarks of cancer: the next generation.," *Cell,* pp. 57-70, 7 January 2000.

[8] G. Cooper, The cell: A molecular Approach. The Development and Causes of Cancer., 2nd edition ed., Sunderland (MA): Sinauer Associates, 2000.

[9] H. Sung, J. Ferlay, R. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal and F. Bray, "Global cancer statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians,* 2021 February 2021.

[10] O. Sanli, J. Dobruch, M. Knowles, M. Burger, M. Alemozaffar, M. Nielsen and Y. Lotan, "Bladder cancer," *Nature,* 13 April 2017.

[11] K. Dietrich, E. Demidenko, A. Schned, M. Zens, J. Heaney and M. Karagas, "Parity, early menopause and the incidence of bladder cancer in women: A case-control study and meta-analysis," pp. 592-599, March 2011.

[12] "Cancer survival by stage at diagnosis for England, 2019," Office for National Statistics.

[13] J. van Oers, C. Adam, S. Denzinger, R. Stoehr, S. Bertz, D. Zaak, C. Stief, F. Hofstaedter, E. Zwarthoff, T. van der Kwast, R. Knuechel and A. Hartmann, "Chromosome 9 deletions are more frequent than FGFR3 mutations in flat urothelial hyperplasias of the bladder," *International Journal of Cancer,* pp. 1212-1215, 5 June 2006.

[14] N. Chow, P. Cairns, C. Eisenberger, M. Schoenberg, D. Taylor, J. Epstein and D. Sidranksy, "Papillary urothelial hyperplasia is a clonal precursor to papillary transitional cell bladder cancer," *International Journal of Cancer,* pp. 514-518, 28 November 2000.

[15] E. Obermann, K. Junker, R. Stoehr, D. Dietmaier, J. Schubert, F. Hofstaedter, R. Knuechel and A. Hartmann, "Frequent genetic alterations in flat urothelial hyperplasias and concomitant

papillary bladder cancer as detected by CGH, LOH, and FISH analyses," *The journal of Pathology,* pp. 50-57, 29 November 2002.

[16] P. Challita-Eid, D. Satpayev, P. Yang, Z. An, K. Morrison, Y. Shostak, A. Raitano, R. Nadell, W. Liu, D. Lortie, L. Capo, A. Verlinsky , M. Leavitt, F. Malik, H. Avina, C. Guevara, N. Dinh, S. Karki, B. Anand, D. Pereira, I. Joseph, F. Donate and D. Stover, "Enfortumab Vedotin Antibody-Drug Conjugate Targeting Nectin-4 Is a Highly Potent Therapeutic Agent in Multiple Preclinical Cancer Models," *Cancer research,* pp. 3003-3013, 15 May 2016.

[17] G. Planes-Laine, P. Rochiqneux, F. Bertucci, A. Chrétien, P. Viens, R. Sabatier and A. Goncalves, "PD-1/PD-L1 Targeting in Breast Cancer: The First Clinical Evidences are Emerging—A Literature Review," *Cancers,* p. 1033, 22 July 2019.

[18] S. Farajzadeh Valilou and N. Rezaei, Vaccines for Cancer Immunotherapy, Academic Press, 2019, pp. 61-74.

[19] P. Bannas, J. Hambach and F. Koch-Nolte, "Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics," *Frontiers in Immunology,* 22 November 2017.

[20] N. Lipman, L. Jackson, L. Trudel and F. Weis-Garcia, "Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources.," *Institute for Laboratory Animal Research Journal,* pp. 258-268, 2005.

[21] T. Olafsen and A. Wu, "Antibody vectors for imaging," *Seminars in Nuclear Medicine Journal,* pp. 167-181, May 2009.

[22] E. Yang and K. Shah, "Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics," *Frontiers in oncology,* 23 July 2020.

[23] S. Steeland, R. van den Broucke and C. Libert, "Nanobodies as therapeutics: big oppurtunities for small antibodies," *Drug Discovery Today*, vol. 21, no. 7, pp. 1076-1113, July 2016.

[24] T. Transue, E. de Genst, M. Ghahroudi, L. Wyns and S. Muyldermans, "Camel single-domain antibody inhibits enzyme by mimicking carbohydrate substrate.," *Proteins,* pp. 515-522, 1 September 1998.

[25] E. de Genst, K. Silence, K. Decanniere, K. Conrath, R. Loris, J. Kinne, S. Muyldermans and L. Wyns, "Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies," *Proceedings of the National Academy of Sciences Journal,* pp. 4586-4591, 21 March 2006.

[26] E. de Genst, K. Silence, K. Decanniere, K. Conrath, R. Loris, J. Kinne, S. Muyldermans and L. Wyns, "Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies," *Proceedings of the National Academy of Sciences USA,* pp. 4586-4591, 13 March 2006.

[27] P. Kunz, K. Zinner, N. Mucke, T. Bartoschik, S. Muyldermans and J. Hoheisel , "The structural basis of nanobody unfolding reversibility and thermoresistance," *Scientific reports,* vol. 8, no. 1, p. 7934, 21 May 2018.

[28] C. Siontorou, "Nanobodies as novel agents for disease diagnosis and therapy," *Internal Journal of Nanomedicine,* vol. 8, pp. 4215-27, 24 July 2013.

[29] E. Ruiz-López and A. Schumacher, "Transportation of Single-Domain Antibodies through the Blood-Brain Barrier," *Biomolecules,* p. 1131, 31 July 2021.

[30] I. van Audenhove and J. Gettemans, "Nanobodies as Versatile Tools to Understand, Diagnose, Visualize and Treat Cancer," *eBioMedicine,* vol. 8, pp. 40-48, June 2016.

[31] S. Smeal, M. Schmitt, R. Pereira, A. Prasad and J. Fisk, "Simulation of the M13 phage life cycle: Assembly of a genetically-structured deterministic chemical kinetic simulation," *Virology,* pp. 259-274, January 2017.

[32] D. Ekiert, R. Friesen, G. Bhabha, T. Kwaks, M. Jongeneelen, W. Yu, C. Ophorst, F. Cox, H. Korse, B. Brandenburg, R. Vogels, J. Brakenhoff, R. Kompier, M. Koldijk, L. Cornelissen, L. Poon, M. Peiris, W. Koudstaal, I. Wilson and J. Goudsmit, "A highly conserved neutralizing epitope on Group 2 Influenza A viruses," *Science,* pp. 843-850, 7 July 2011.

[33] D. Ekeirt, A. Kashyap, J. Steel, A. Rubrum, G. Bhabha, R. Khayat, J. Lee, M. Dillon, R. O'Neil, A. Faynboym, M. Horowitz, L. Horowitz, A. Ward, P. Palese, R. Webby, R. Lerner, R. Bhatt and I. Wilson, "Cross-neutralization of influenza A viruses mediated by single antibody loop," *Nature,* pp. 526-532, 16 September 2012.

[34] H. Karauzum, G. Chen, L. Abaandou, M. Mahmoudieh, A. Boroun, S. Shulenin, V. Devi, E. Stavale, K. Warfield, L. Zeitlin, C. Roy, S. Sidhu and M. Aman, "Synthetic Human Monoclonal Antibodies toward Staphylococcal Enterotoxin B (SEB) Protective against Toxic Shock Syndrome," *Journal of biological chemistry,* pp. 25203-25215, 20 July 2012.

[35] S. Koide and S. Sidhu, "The Importance of Being Tyrosine: Lessons in Molecular Recognition from Minimalist Synthetic Binding Proteins," *ACS chemical biology,* pp. 325-334, 19 March 2009.

[36] G. Smith, "Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface," *Science,* pp. 1315-1317, 14 June 1985.

[37] L. Ledsgaard, M. Kilstrup, A. Karatt-Vellatt, J. McCafferty and A. Laustsen, "Basics of antibody phage display technology," *Toxins,* 9 June 2018.

[38] K. Chatalic, J. Veldhoven-Zweistra, M. Bolkestein, S. Hoeben, G. Koning, O. Boerman, M. de Jong and W. van Weerden, "A Novel ^{111}In-Labeled Anti-Prostate-Specific Membrane Antigen Nanobody for Targeted SPECT/CT Imaging of Prostate Cancer," *Journal of Nuclear medicine,* vol. 56, no. 7, pp. 1094-1099, 14 May 2015.

[39] T. Kaprio, J. Hagstrom, L. Andersson and C. Haglund, "Tetraspanin CD63 independently predicts poor prognosis in colorectal cancer," *Histology and histopathology,* vol. 35, no. 8, pp. 887-892, 19 February 2020.

[40] R. Reyes, B. Cardenes, Y. Machado-Pineda and C. Cabanas, "Tetraspanin CD9: A Key Regulator of Cell Adhesion in the Immune System," *Frontiers in Immunology,* vol. 9, 18 April 2018.

[41] H. Romanska and F. Berditchevski, "Tetraspanins in human epithelial malignancies," *Journal of pathology,* vol. 223, no. 1, pp. 4-14, 2011.

[42] J. Hyeong, S. Min, M. Park, C. Lee, J. Park, J. Chae and K. Moon, “Expression of CD9 and CD82 in clear cell renal cell carcinoma and its clinical significance,” *Pathology - research and practice,* vol. 210, no. 5, pp. 285-290, May 2014.

[43] J. Garner, M. Herr, K. Hodges and L. Jennings, “The utility of tetraspanin CD9 as a biomarker for metastatic clear cell renal cell carcinoma,” *Biochemical and Biophysical Research Communications,* vol. 471, no. 1, pp. 21-25, 26 February 2016.

[44] J. Kenneth and J. Clemetson, “Platelet Receptors,” in *Platelets,* 3 ed., Academic Press, 2013, pp. 169-194.

[45] M. Nishibori, B. Cham, A. McNicol, A. Shalev, N. Jain and J. Gerrard, “The protein CD63 is in platelet dense granules, is deficient in a patient with Hermansky-Pudlak syndrome, and appears identical to granulophysin,” *The Journal of Clinical Investigation,* pp. 1775-1782, April 1993.

[46] A. Lupia, S. Peppicelli, E. Witort, F. Bianchini, V. Carloni, N. Pimpinelli, C. Urso, L. Borgognoni, S. Capaccioli, L. Calorini and M. Lulli, “CD63 tetraspanin is a negative driver of epithelial-to-mesenchymal transition in human melanoma cells,” *Journal of investigative dermatology,* vol. 134, no. 12, pp. 2947-2956, Dec 2014.

[47] H. Hotta, A. Ross, K. Huebner, M. Isobe, S. Wendeborn, M. Chao, R. Ricciardi, Y. Tsujimoto, C. Croce and H. Koprowski, “Molecular cloning and charactirization of an antigen associated with early stages of melanoma tumor progression.,” *Cancer research,* vol. 48, no. 11, pp. 2955-2962, 1 June 1988.

[48] B. Atkinson, C. Ernst, B. Ghrist, M. Herlyn, M. Blaszczyk, A. Ross, D. Herlyn, Z. Steplewski and H. Koprowski, “Identification of melanoma-associated antigens using fixed tissue screening of antibodies,” *Cancer research,* vol. 44, no. 6, pp. 2577-2581, June 1984.

[49] P. Chames, “CESAME Manual Version 2.0,” 1996.

[50] “LabNed Plasmid Midiprep Kit - LN2400004,” 2017. [Online]. Available: https://api.labned.com/img/pdf/24/LN2400004.pdf.

[51] “WizardR SV Gel and PCR Clean-Up System Protocol,” (n.d.). [Online]. Available: https://nld.promega.com/resources/protocols/technical-bulletins/101/wizard-sv-gel-and-pcr-cleanup-system-protocol/. [Accessed 21 June 2022].

[52] M. Green and J. Sambrook, “The Inoue Method for Preparation and Transformation of Competenent Escherichia coli: "Ultracompetent" Cells.,” *Cold Spring Harbor Protocols,* vol. 2020, no. 6, p. 101196, 1 June 2020.

[53] J. Zhang, K. Kobert, T. Flouri and A. Stamatakis, “PEAR: a fast and accurate Illumina Paired-End reAd mergeR.,” *Bioinformatics,* pp. 614-620, 1 March 2014.

[54] S. Andrews, “FastQC: A Quality Control Tool for High Throughput Sequence Data.,” 2010. [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

[55] M. Martin, “Cutadapt removes adapter sequences from high-throughput sequencing reads.,” *EMBnet.journal,* pp. 10-12, May 2011.

[56] D. Bolotin, S. Poslavsky and L. Mitrophanov, “MiXCR: software for comprehensive adaptive immunity profiling.,” *Nature Methods,* pp. 380-381, 29 April 2015.

[57] K. L. Chatalic, J. Veldhoven-Zweistra, M. Bolkestein, S. Hoeben, G. A. Koning, O. C. Boerman, M. de Jong and W. M. van Weerden, "A Novel ^{111}In-Labeled Anti-Prostate-Specific Membrane Antigen Nanobody for Targeted SPECT/CT Imaging of Prostate Cancer," *Journal of Nuclear Medicine,* vol. 56, no. 7, pp. 1094-1099, 14 May 2015.

[58] G. Dotto, K. Horiuchi and N. Zinder, "The functional origin of bacteriophage f1 DNA replciation: Its signals and domains," *Journal of Molecular Biology,* pp. 507-521, 5 February 1984.

[59] N. Heynickx, K. Herrmann, K. Vermeulen, S. Baatout and A. Aerts, "The salivary glands as a dose limiting organ of PSMA- targeted radionuclide therapy: A review of the lessons learnt so far.," *Nuclear Medicine and Biology,* pp. 30-39, July-August 2021.

[60] L. Chen, S. Zou, D. Li, J. Zhou, Z. Cheng, J. Zhao, Y. Zhu, D. Kuang and X. Zhu, "Prostate-specific membrane antigen expression in hepatocellular carcinoma, cholangiocarcinoma, and liver cirrhosis," *World Journal of Gastroenterology,* pp. 7664-7678, 28 Dec 2020.

[61] D. Jiao, Y. Li, D. Han, F. Yang, J. Wu, S. Shi, F. Tian, Z. Guo, W. Xi, G. Li, A. Zhao, A. Yang, W. Qin, H. Wang and W. Wen, "Expression of Prostate-specific Membrane Antigen in Tumor-Associated Vasculature Predicts Poor Prognosis in Hepatocellular Carcinoma.," *Clinal and Translational Gastroenterology,* pp. 1-7, 1 May 2019.

[62] J. Carpenter and J. Crowe, "Modes of stabilization of a protein by organic solutes during desiccation," *Cryobiology,* pp. 459-470, October 1988.

[63] D. Nusinow, J. Szpyt, M. Ghandi, C. Rose, E. McDonald, M. Kalocsay, J. Jané-Valbuena, E. Gelfand, D. Schweppe, M. Jedrychowski, J. Golji, D. Porter, T. Rejtar, Y. Wang, G. Kryukov, F. Stegmeier, B. Erickson, L. Garraway, W. Sellers and S. Gygi, "Quantitative Proteomics of the Cancer Cell Line Encyclopedia.," *Cell,* pp. 387-402, 23 Januari 2020`.

[64] S. Magdeldin and T. Yamamoto , "Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues," *Proteomics,* pp. 1045-1058, April 2012.

[65] D. Hingorani, M. Allevato, M. Camargo, J. Lesperance, M. Quraishi, J. Aguilera, I. Franiak-Pietryga, D. Scanderbeg, Z. Wang, A. Molinolo, D. Alvarado, A. Sharabi, J. Bui, E. Cohen, S. Adams, J. Gutkind and S. Advani, "Monomethyl auristatin antibody and peptide drug conjugates for trimodal cancer chemo-radio-immunotherapy," *Nature Communications,* July 05 2022.

Supplementary Data

1 Major research project portfolio

	Institution	Workload (ECTS)
General courses		
Plasma medicine	Utrecht University	3 EC
Biology of disease: Immunity and Infection	Utrecht University	3 EC
Orientation on Presentation and Career	Utrecht University	2.5 EC
Navigation towards Personal Excellence	Utrecht University	1.0 EC
Seminars and workshops		
JNI oncology lectures	Erasmus MC	1.5 EC
Department Journal Club	Erasmus MC	0.5 EC
Symposium Academic Centre of Excellence 'Tumor Immunology and Immune Therapy'	Erasmus MC	0.4 EC
SUMMA symposium 'Technological advances in healthcare; how far can we go and how far should we go?'	Utrecht University	0.3 EC
Life Science Academy seminars	Utrecht University	0.3 EC
'A dialogue on animals as organ donor for humans'		
'The promise of open science'		
'Gendered innovation in biomedicine and health'		
'Life science and society – concept of planetary health'		
'True passive patient monitoring to enable better health care'		
'Chromosome replication in Genome and Epigenome Maintenance'		
QBio Symposium 'Age is just a number'!	Utrecht University	0.2 EC
Presentations		
Oral presentation scientific lab meetings, three times	Erasmus MC	0.3 EC

2 Barcode sequences used for PCR

Barcode	Sequence	Forward (F) or Reverse (R)
1	CTAAACTACGG	F
2	GCAGATCCAAC	F
3	CATCACATAGG	F
4	TGGTATGGGAGA	F
5	CTTTAAGGGTGA	F
6	AGCAACATCCTA	F
7	GTTGCGTTTCTTC	F
8	TGTCCGACCAATC	F
9	GGTACGCAATTTC	F
10	CAGCCACCCATCTA	F
11	GTCTCGCAAGCCTA	F
12	AGGAGTAAAGCCTA	F
B	AACCACGTAAC	R
C	CTCCCATACCAC	R
D	ACACCCTTAAAG	R
E	ACAGAAACGCAAC	R
F	ACTTGGTCGGACA	R
G	TCGATGGGTGGCTG	R
H	TCGGCTTGCGAGAC	R

Table 7 | Barcode combinations of the forward and reverse primer sequences.

3 Macro RGB values

```
dir = "directory”/;
list = getFileList(dir);
for (i = 0; i < list.length; i++){
        open(dir+list[i]);
        path = dir + list[i];
        x = File.getName(path)+ "/";
        name = File.getName(path);
        name = substring(name, 0, lastIndexOf(name, "."));

        //The image is splitted into three channels (R,G,B)
        run("Make Composite");
        run("Set Measurements...","mean redirect = None decimal=0");
        Stack.setChannel("1");
        run("Measure");
        Stack.setChannel("2");
        run("Measure");
        Stack.setChannel("3");
        run("Measure");
        close();

        // The mean grey value of the three channels is measured
        R = getResult("Mean",0);
        G = getResult("Mean",1);
        B = getResult("Mean",2);

//The output is printed and ready for further analysis in excel
print(name + "," + d2s(R,0) + "," + d2s(G,0) + "," + d2s(B,0));
IJ.deleteRows(0,2);
}
```

4 FroCellCA H6, Nobi, A7

Figure 34 – 36 | Images of DAB staining showing the affinity between Nbs (5 x 10^{10} ø/mL) and twenty different cell lines. For immunocytochemical staining, we included six BlCa cell lines (T24, TCCSUP, RT112, J82, JON, 5637), seven PCa cell lines (22Rv1, DU145, LNCaP, MDAPCa2b, PC3, PC346C, VCaP) and other Ca cell lines (COLO205, H460, HCC1806, HEK293, Hep3B, HUVEC and MCF7). Images were taken at 20x magnification using the BX41TF Olympus microscope.

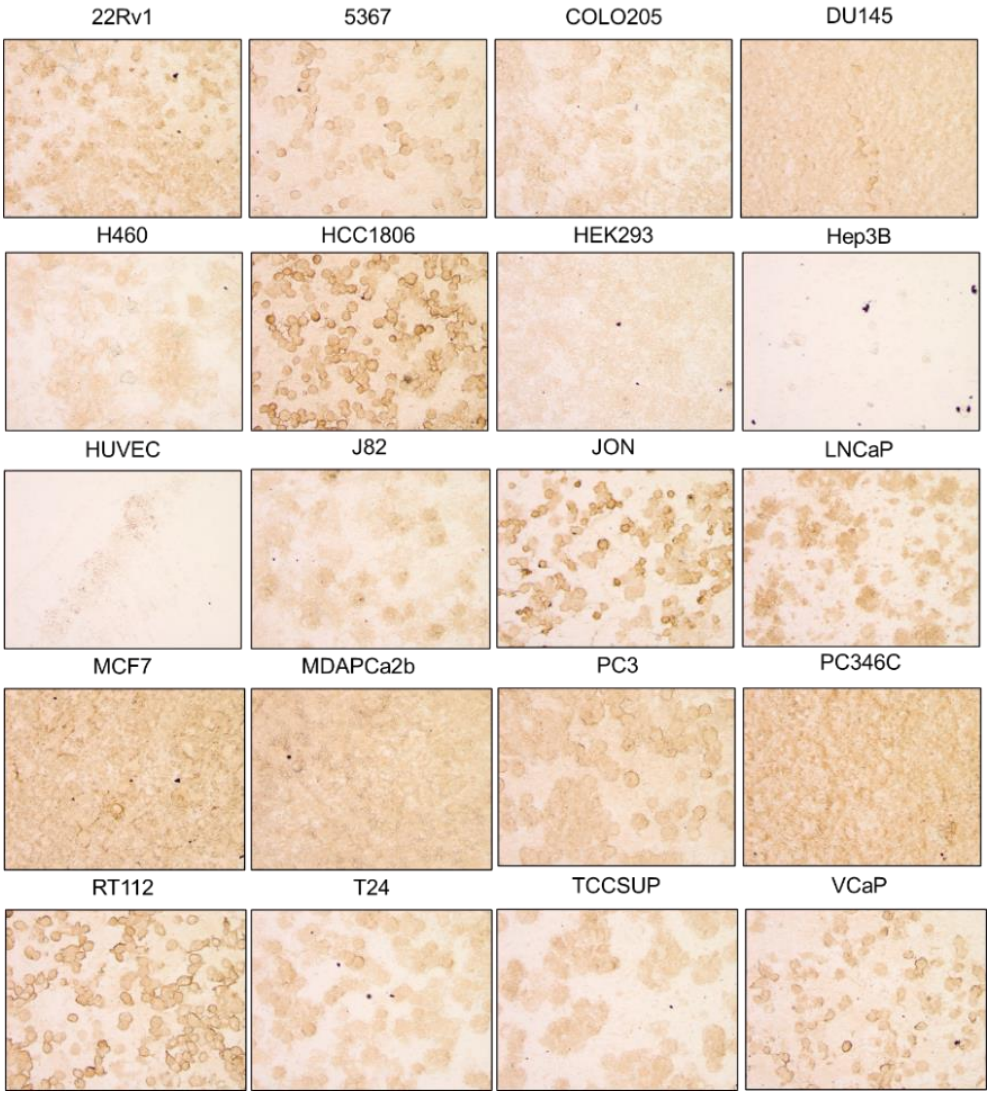


Figure 34 | Positive control H6 Nb binding to different Ca cell lines.

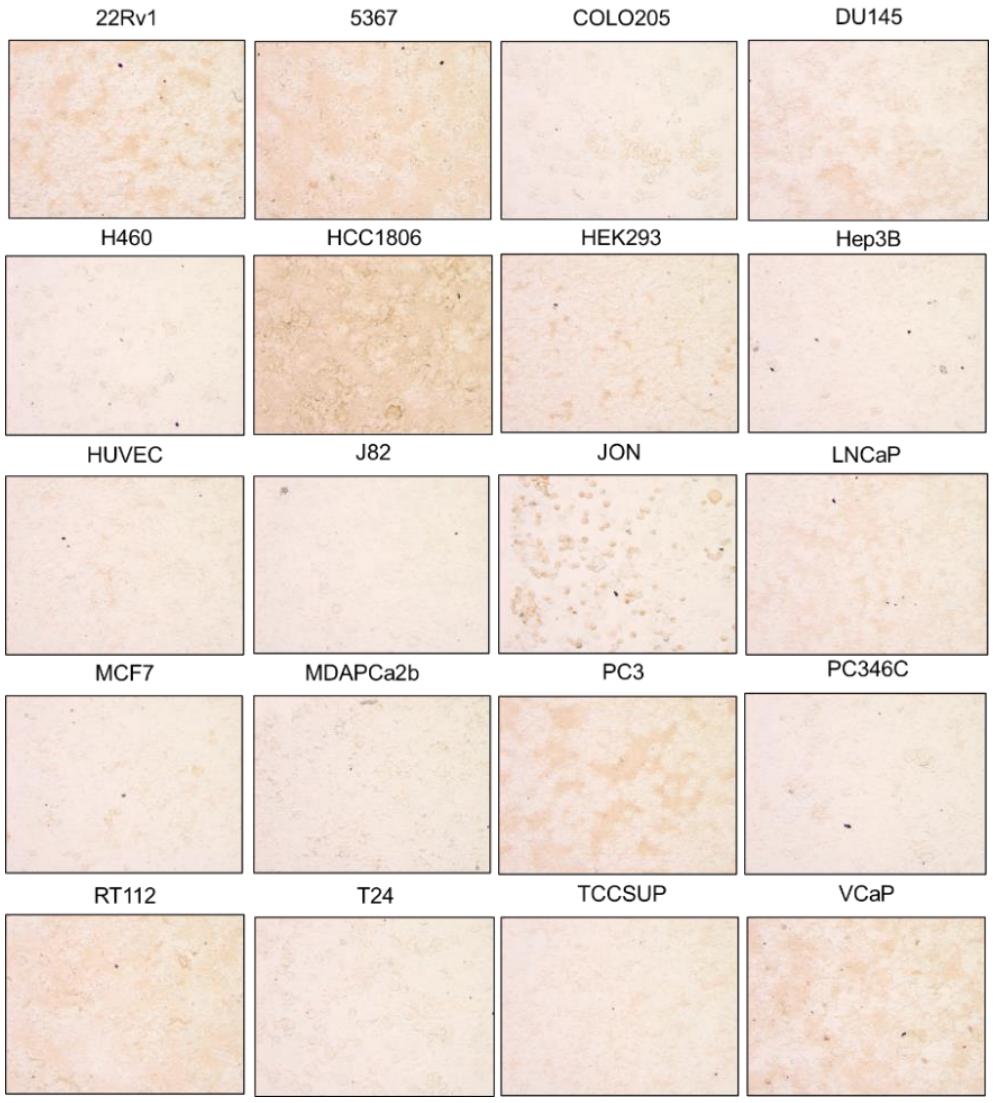


Figure 35 | Negative control NoBi Nb binding to different Ca cell lines.

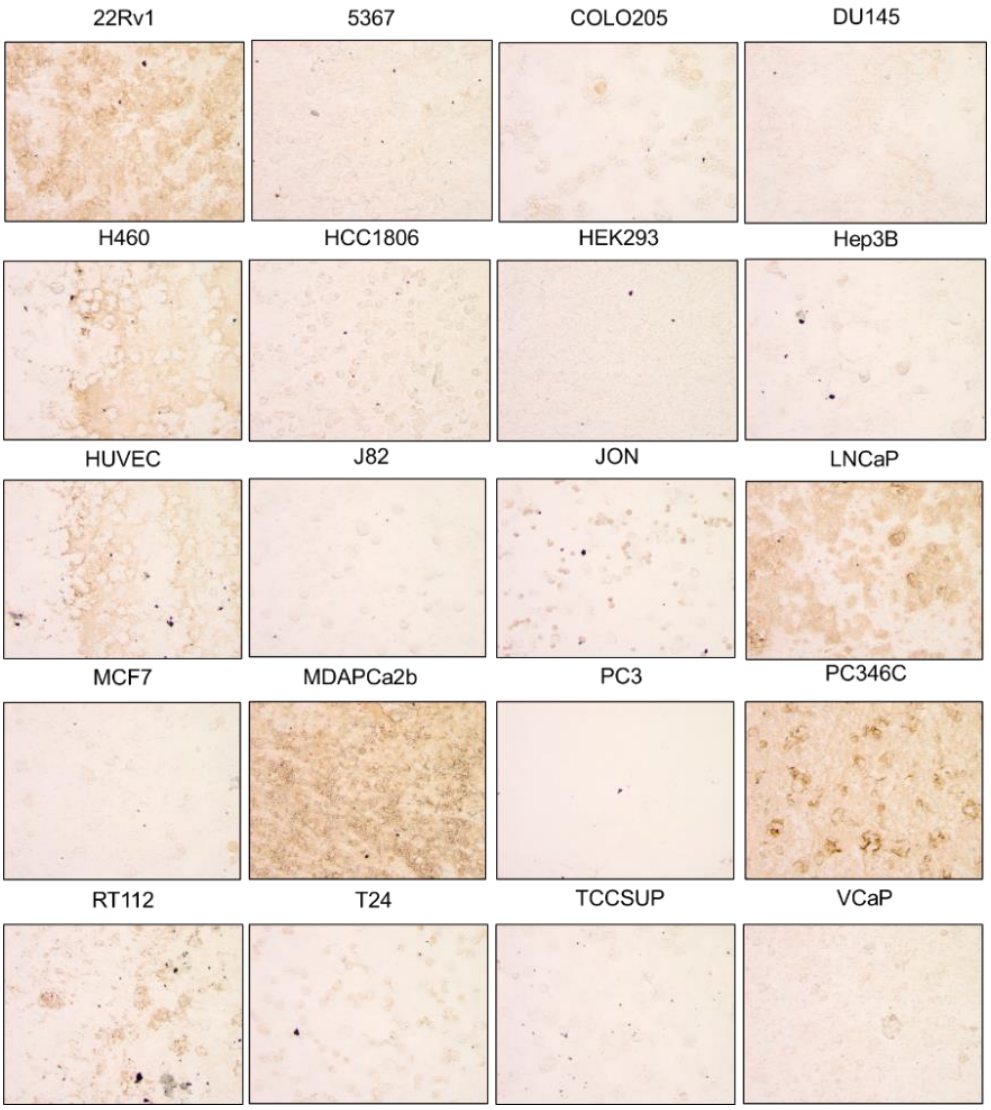


Figure 36 | PSMA-specific binding of A7 Nb binding to different Ca cell lines.

5 FroCellCA pre-selected libraries

Figure 37 – 39 | Images of DAB staining showing the affinity between two different M13-phage libraries and twenty different cell lines. All preBlC-libraries were pre-selected for six BlCa cell lines: 5637, T24, TCCSUP, JON, J82 and RT112. For immunocytochemical staining, we included six BlCa cell lines (T24, TCCSUP, RT112, J82, JON, 5637), seven PCa cell lines (22Rv1, DU145, LNCaP, MDAPCa2b, PC3, PC346C, VCaP) and other Ca cell lines (COLO205, H460, HCC1806, HEK293, Hep3B, HUVEC and MCF7).

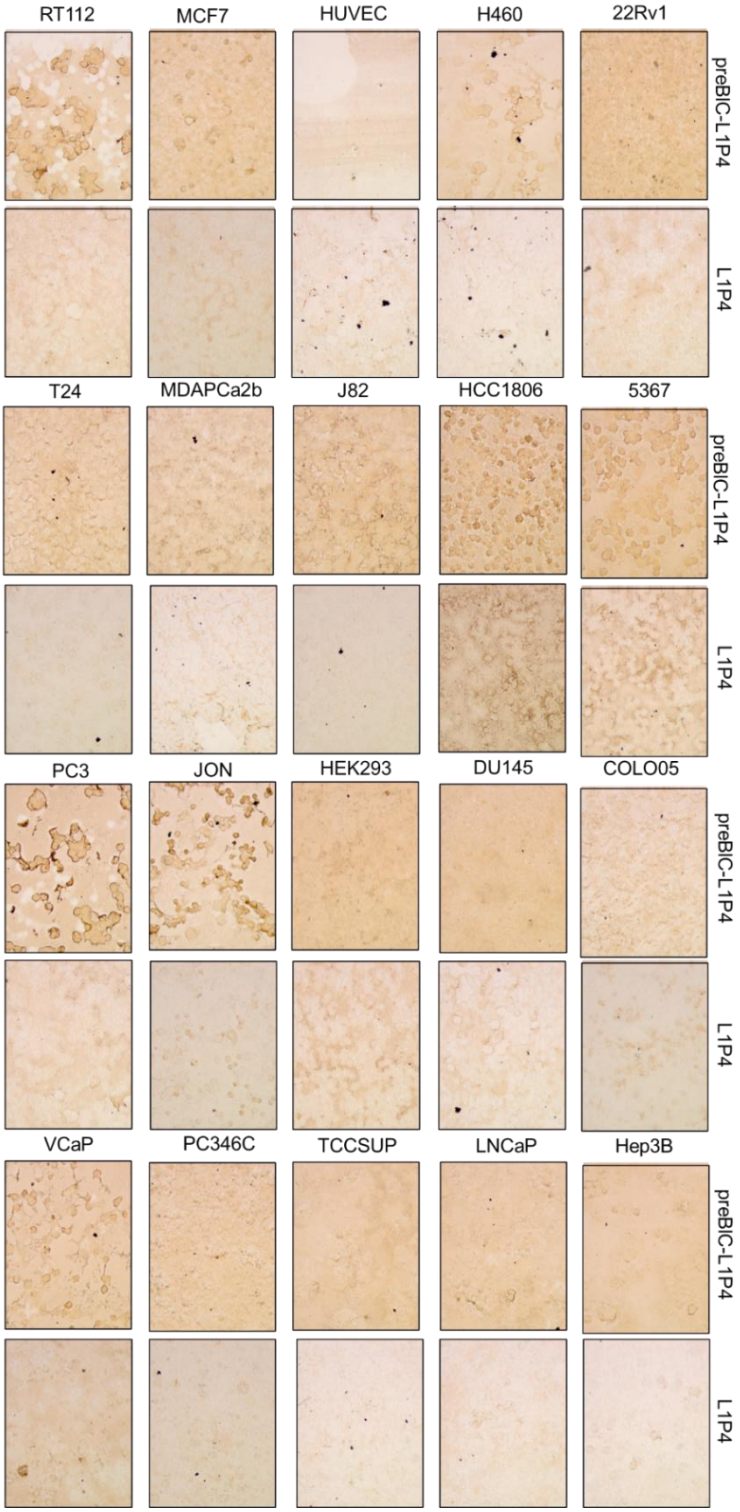


Figure 37 | preBlC-L1P4 versus L1P4 library binding to different Ca cell lines.

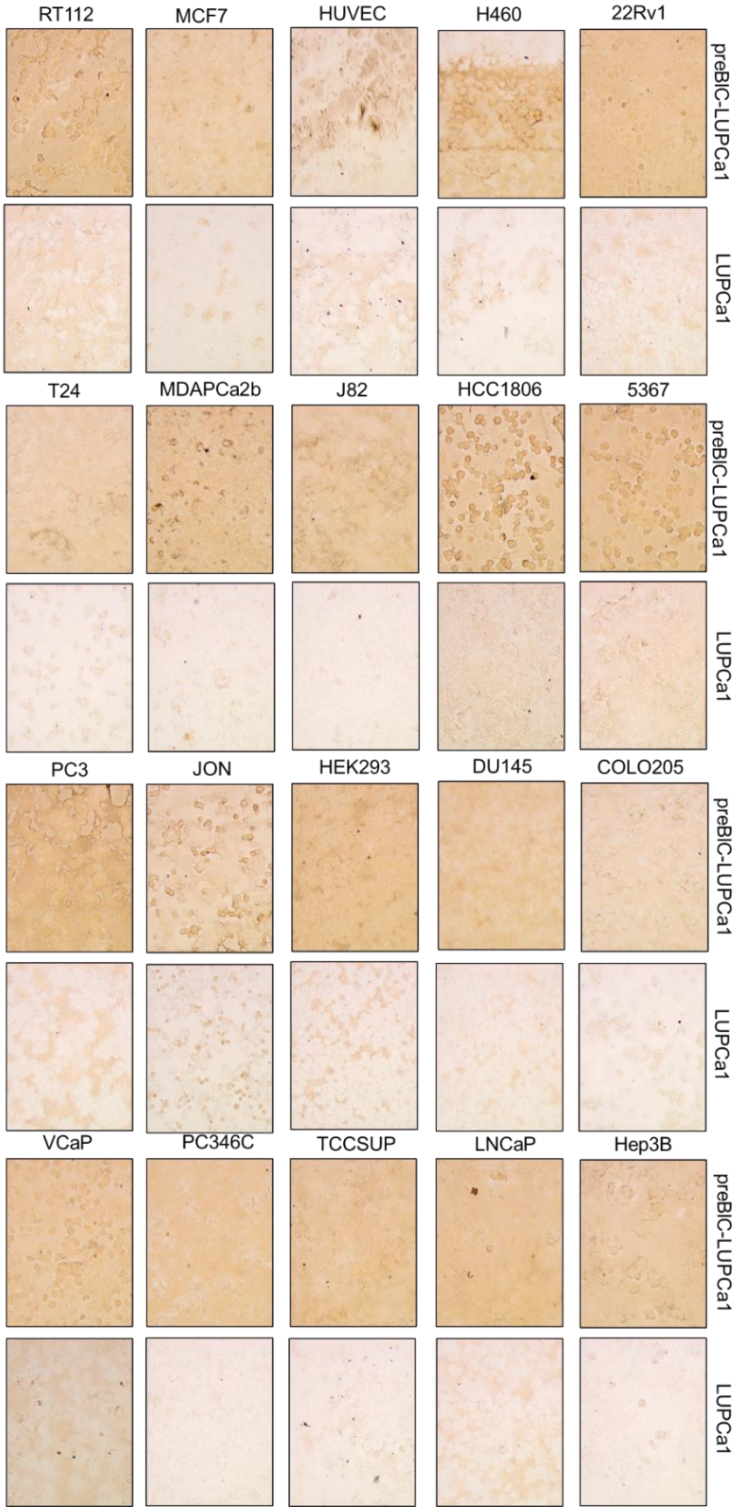


Figure 38 | preBlC-LUPCa1 versus LUPCa1 library binding to different Ca cell lines.

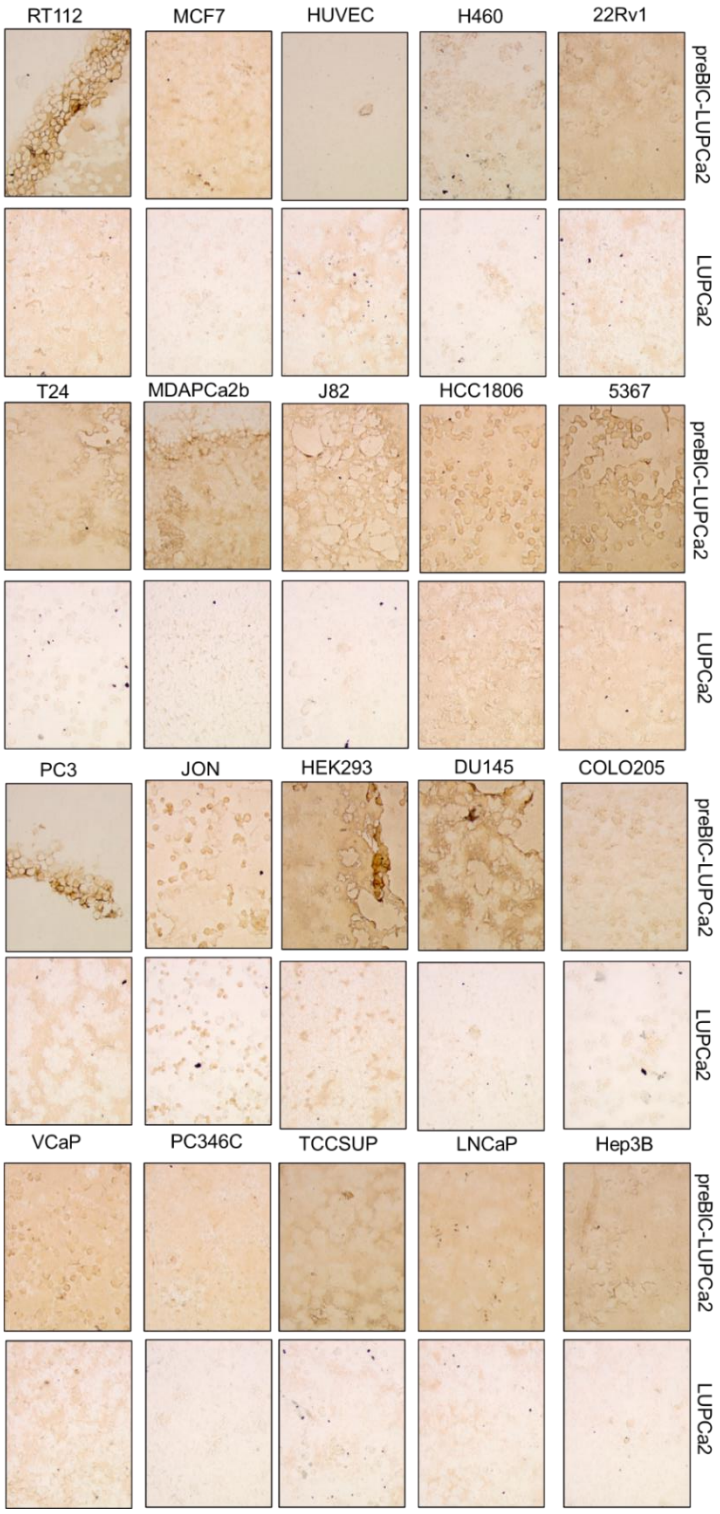


Figure 39 | preBlC-LUPCa2 versus LUPCa2 library binding to different Ca cell lines.

6 FroCellCA BlCaNbs

Figure 40 – 54 | Images of DAB staining showing the affinity between a specific Nb (5 x 10^{10} ø/mL) and twenty different cell lines. Images were taken at 20x magnification using the BX41TF Olympus microscope. For immunocytochemical staining, we included six BlCa cell lines (T24, TCCSUP, RT112, J82, JON, 5637), seven PCa cell lines (22Rv1, DU145, LNCaP, MDAPCa2b, PC3, PC346C, VCaP) and other Ca cell lines (COLO205, H460, HCC1806, HEK293, Hep3B, HUVEC and MCF7).

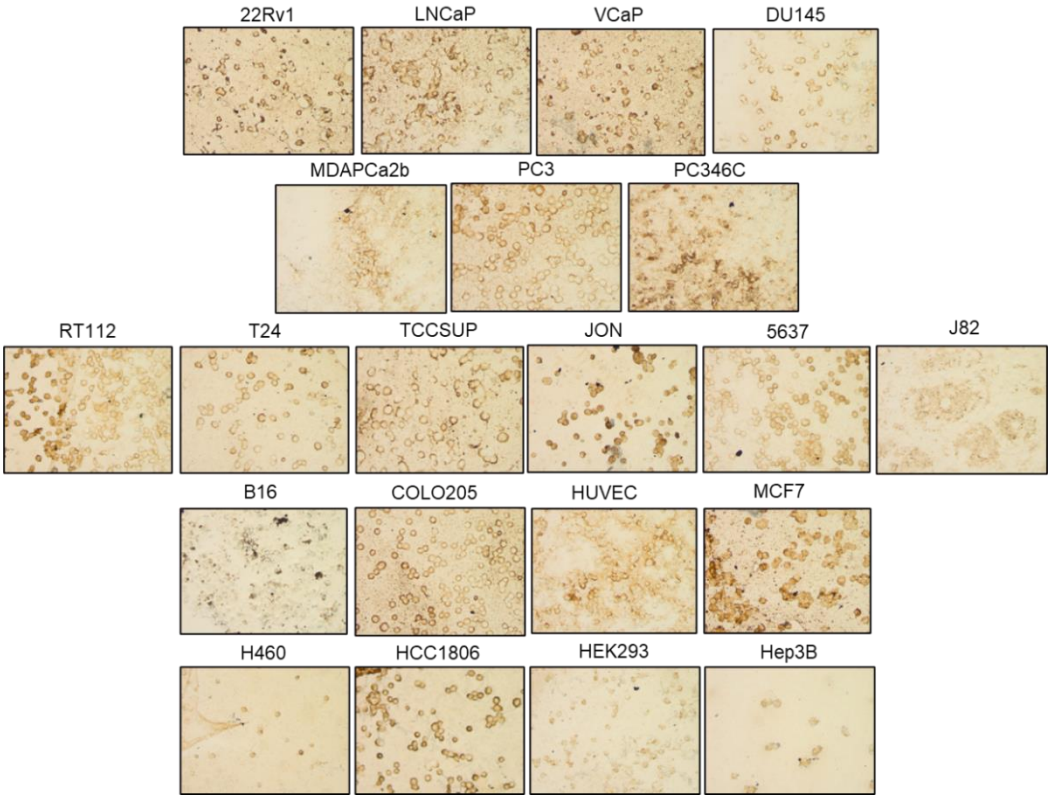


Figure 40 | H6 Nb binding to different Ca cell lines.

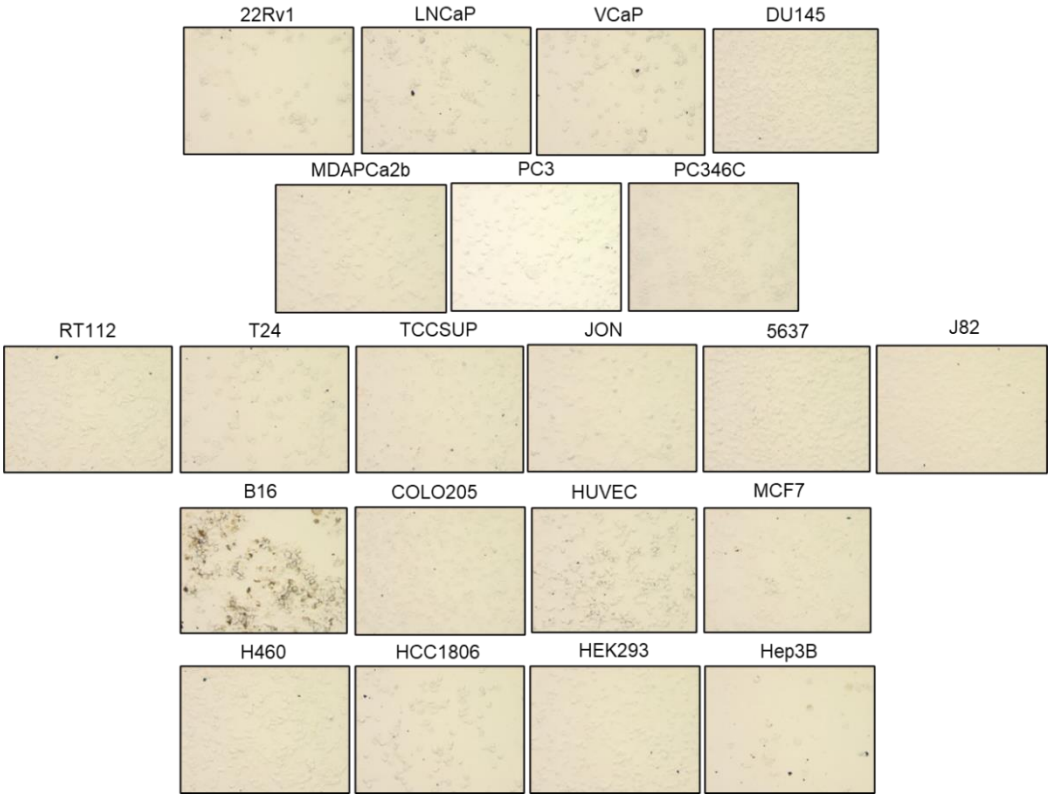


Figure 41 | A7 Nb binding to different Ca cell lines.

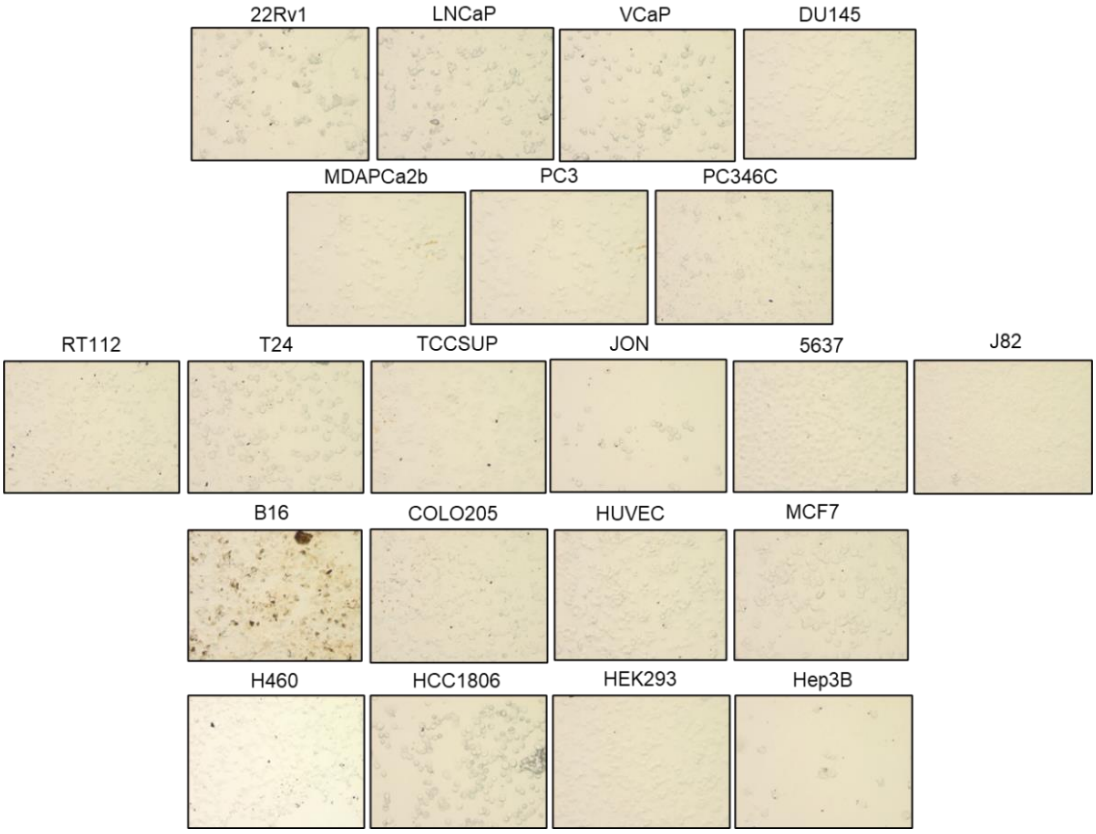


Figure 42 | NoBi Nb binding to different Ca cell lines.

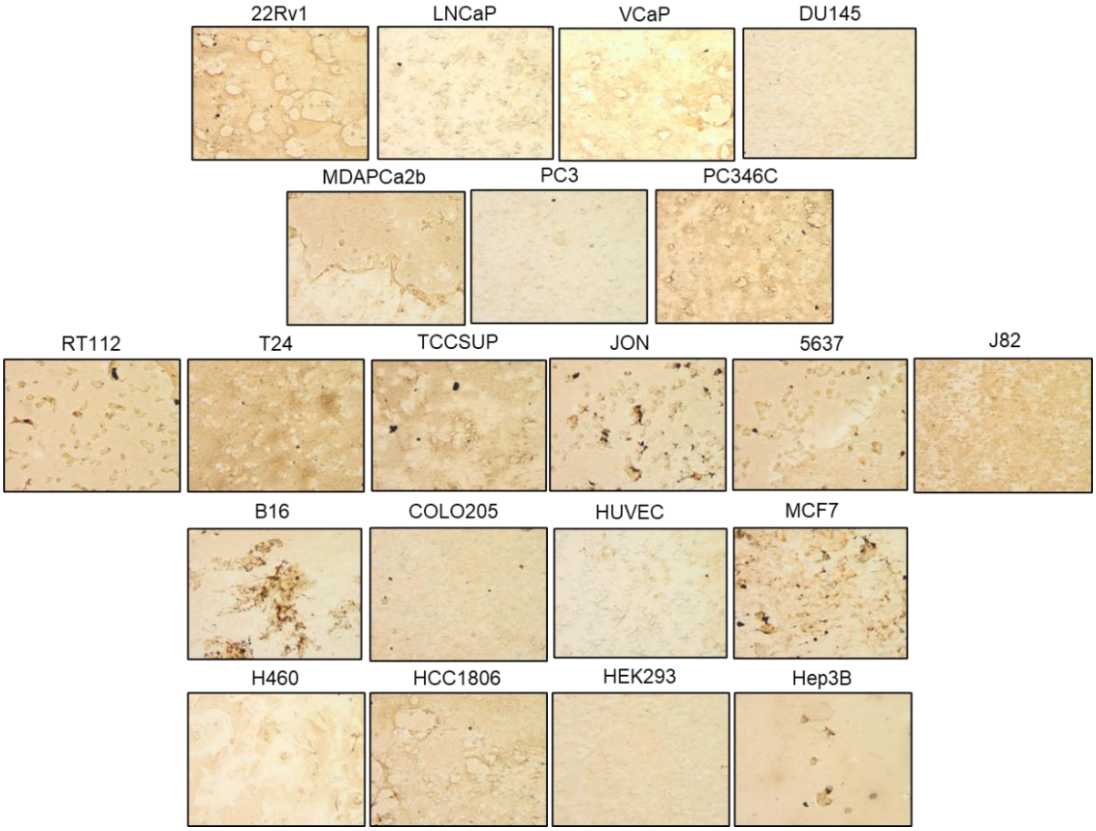


Figure 43 | BlCaNb35 Nb binding to different Ca cell lines

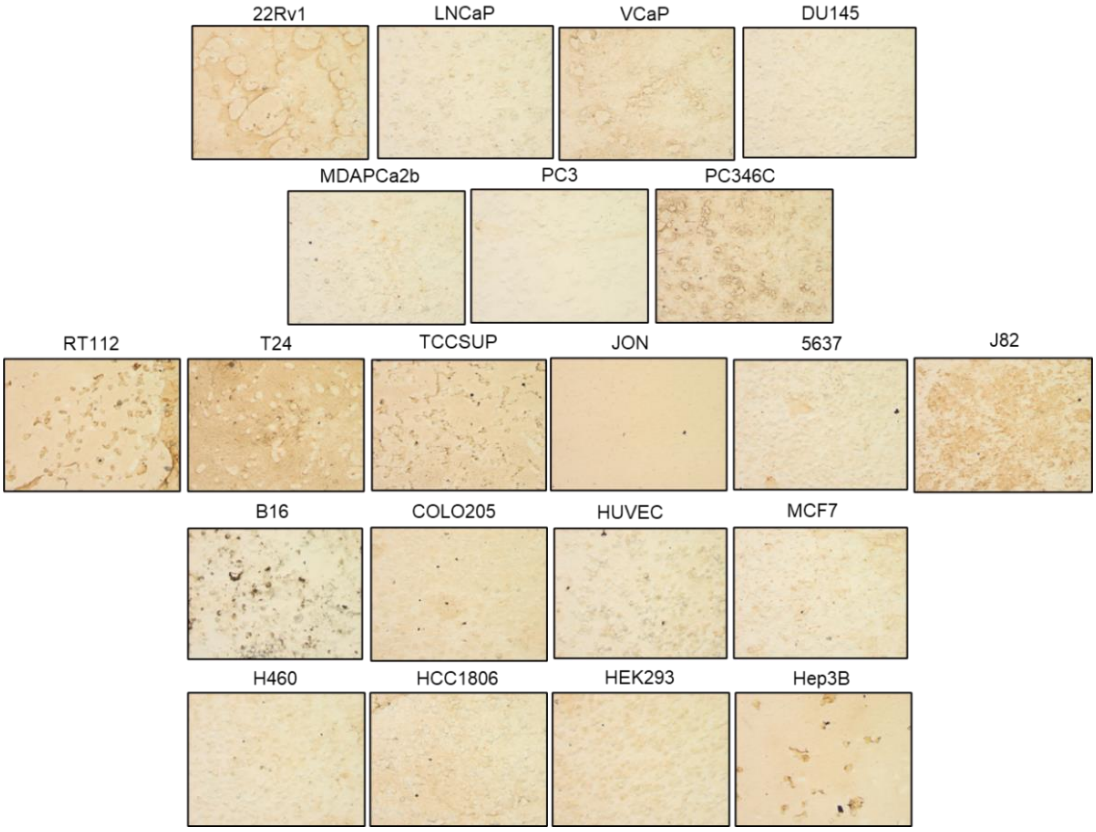


Figure 44 | BlCaNb36 Nb binding to different Ca cell lines.

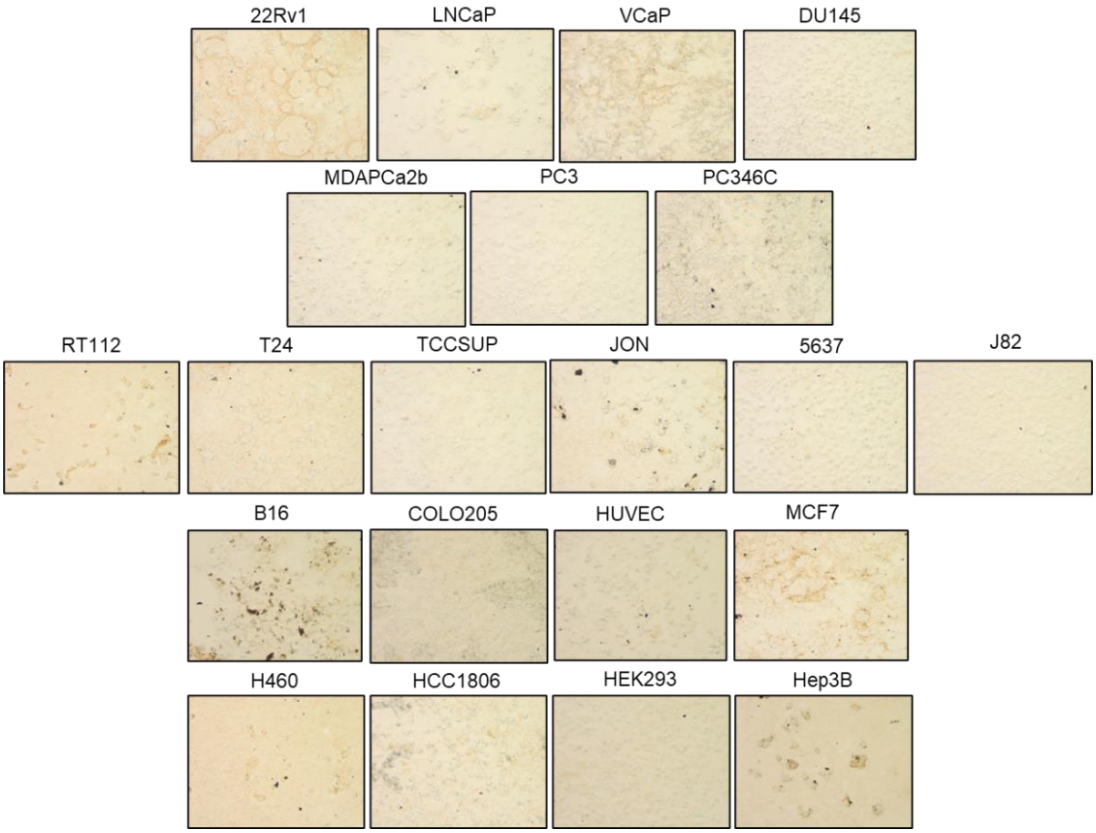


Figure 45 | BlCaNb37 Nb binding to different Ca cell lines.

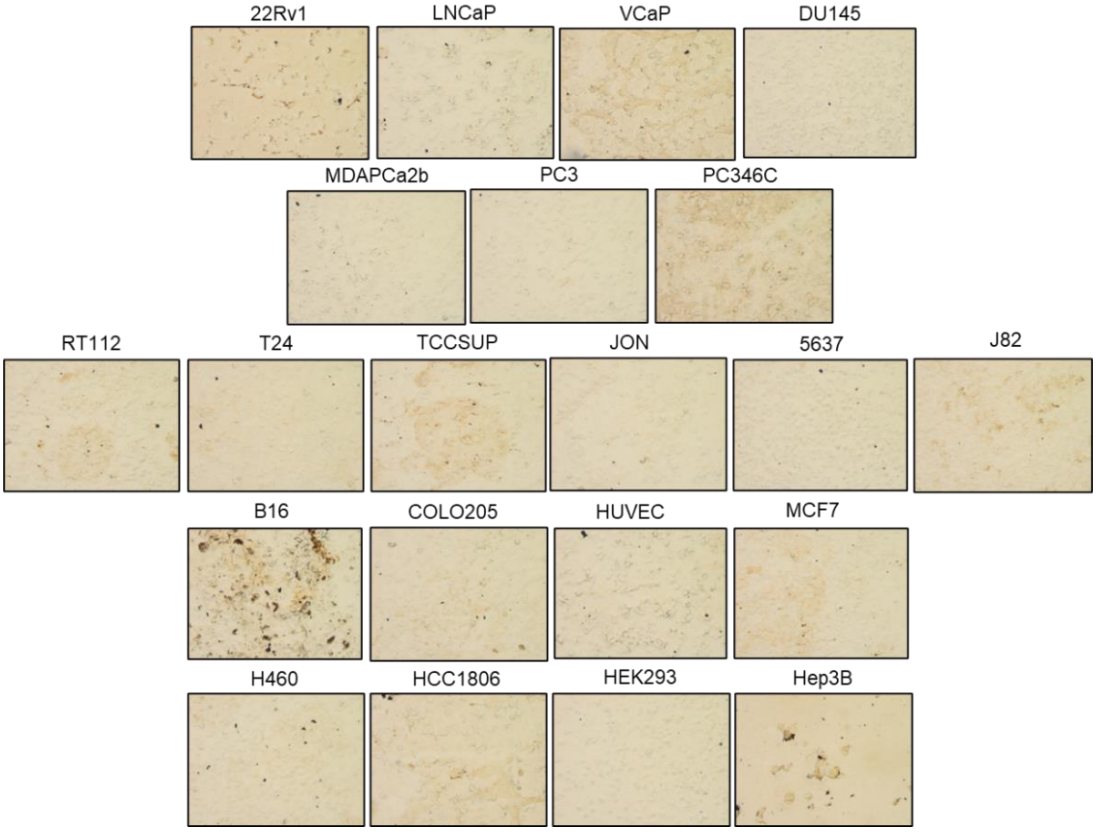


Figure 46 | BlCaNb38 Nb binding to different Ca cell lines.

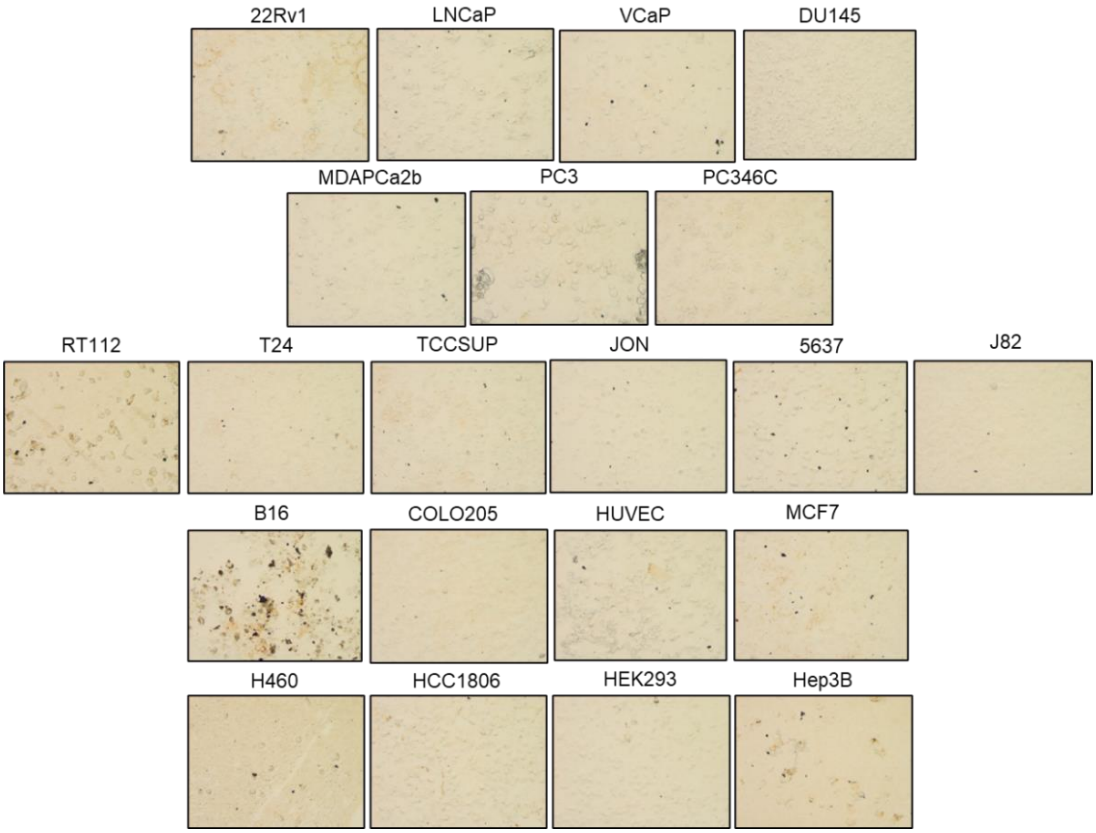


Figure 47 | BlCaNb39 Nb binding to different Ca cell lines.

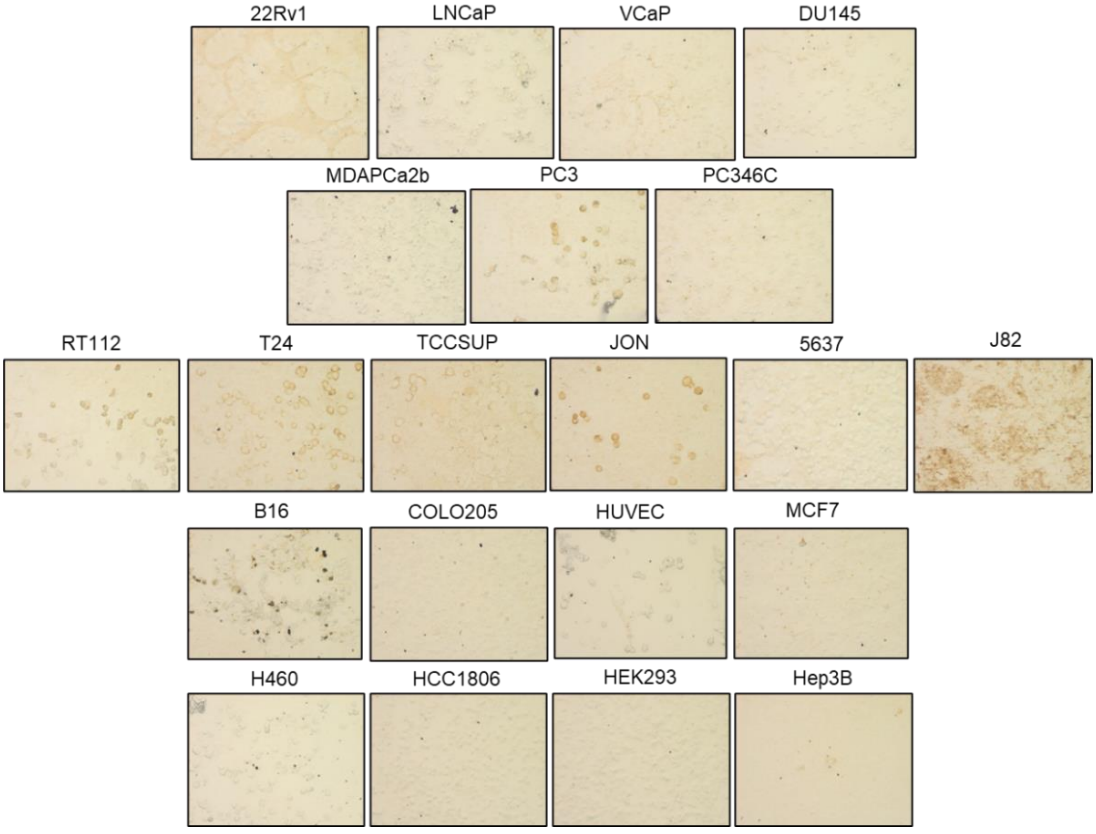


Figure 48 | BlCaNb40 Nb binding to different Ca cell lines.

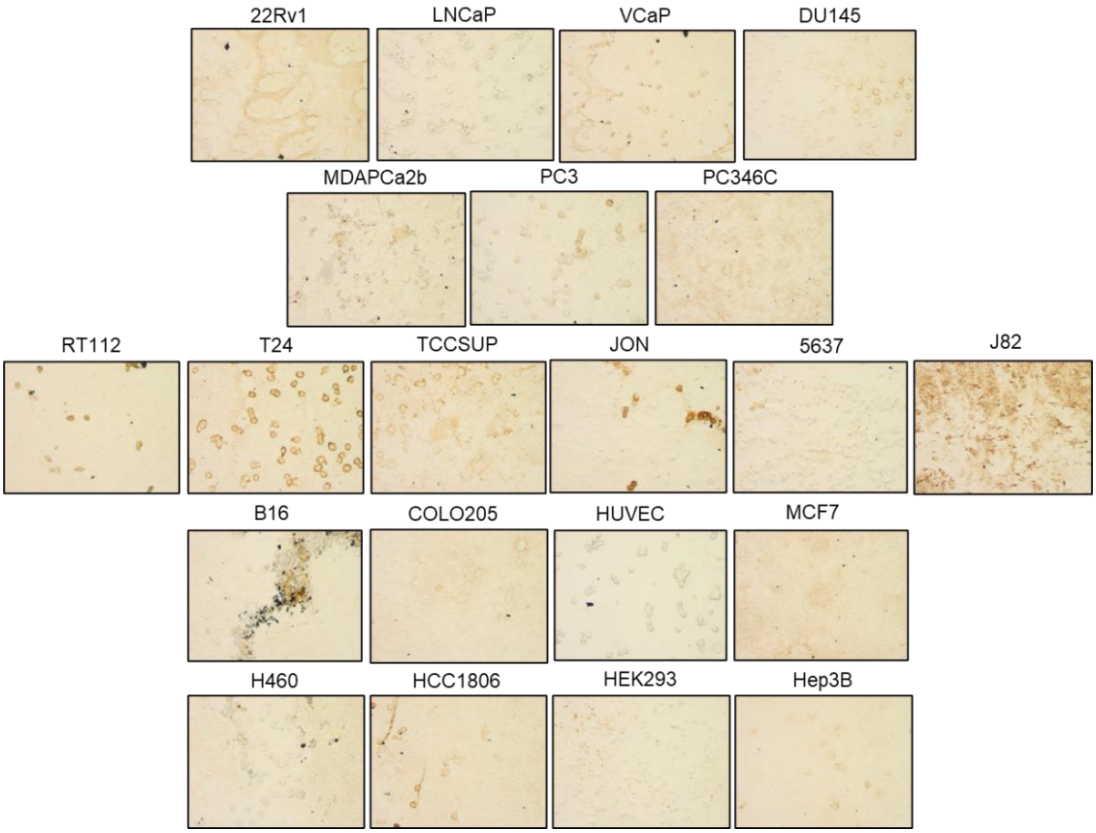


Figure 49 | BlCaNb41 Nb binding to different Ca cell lines.

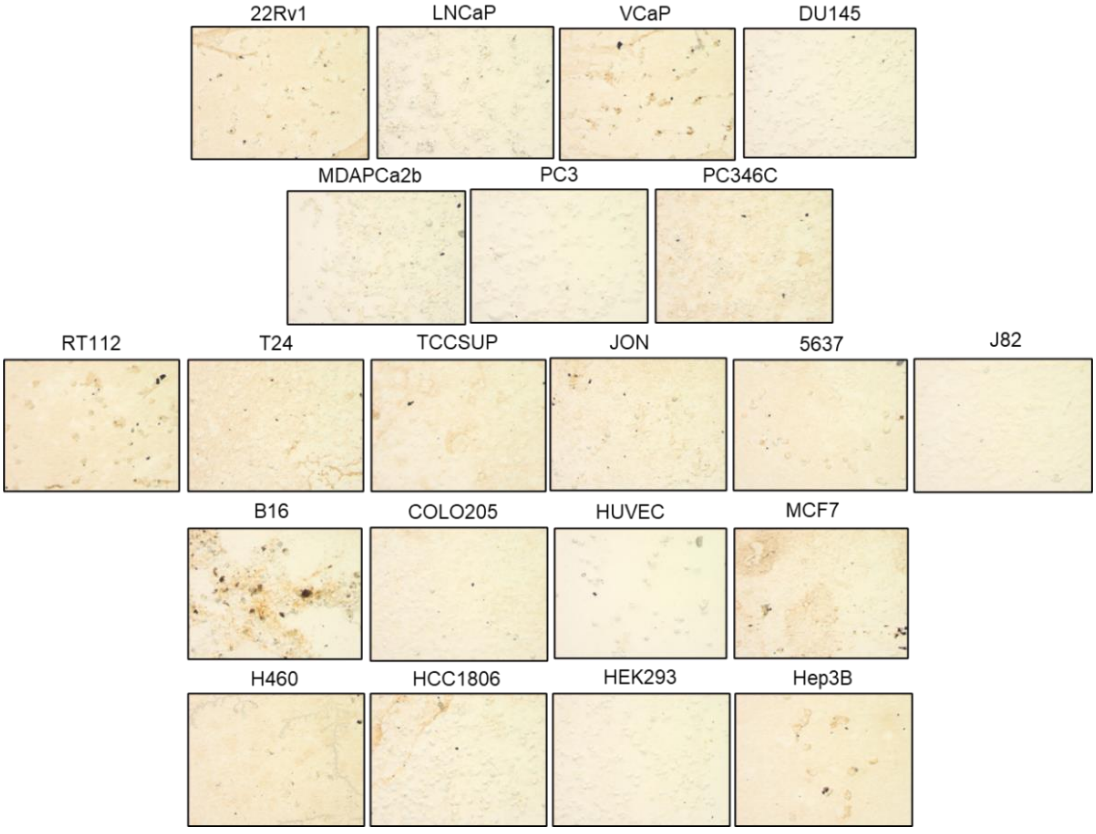


Figure 50 | BlCaNb42 Nb binding to different Ca cell lines.

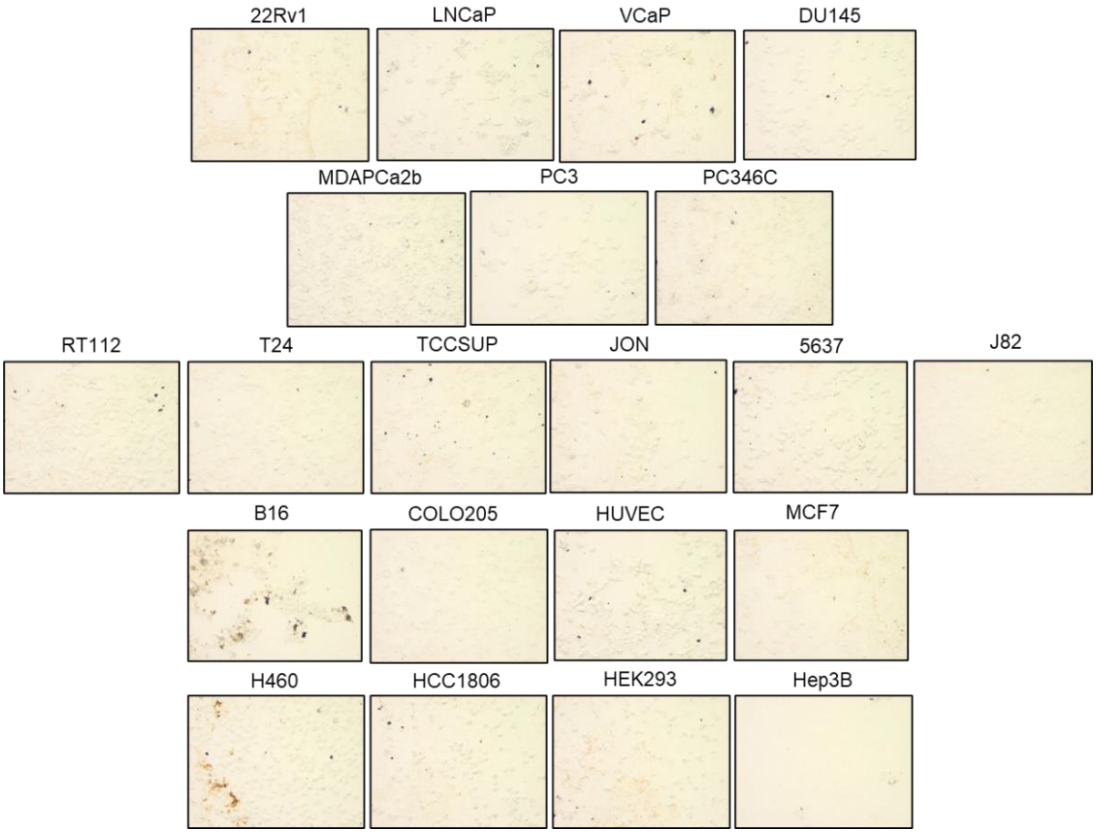


Figure 51 | BlCaNb43 Nb binding to different Ca cell lines.

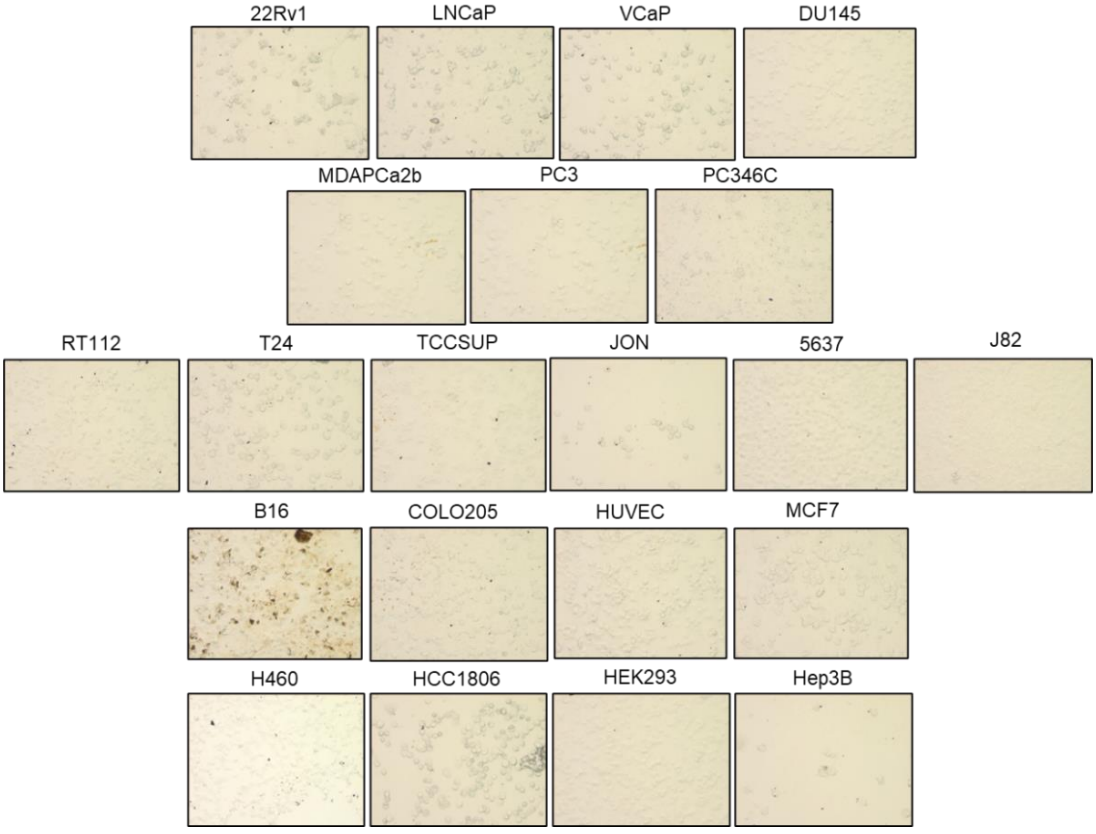


Figure 52 | BlCaNb44 Nb binding to different Ca cell lines.

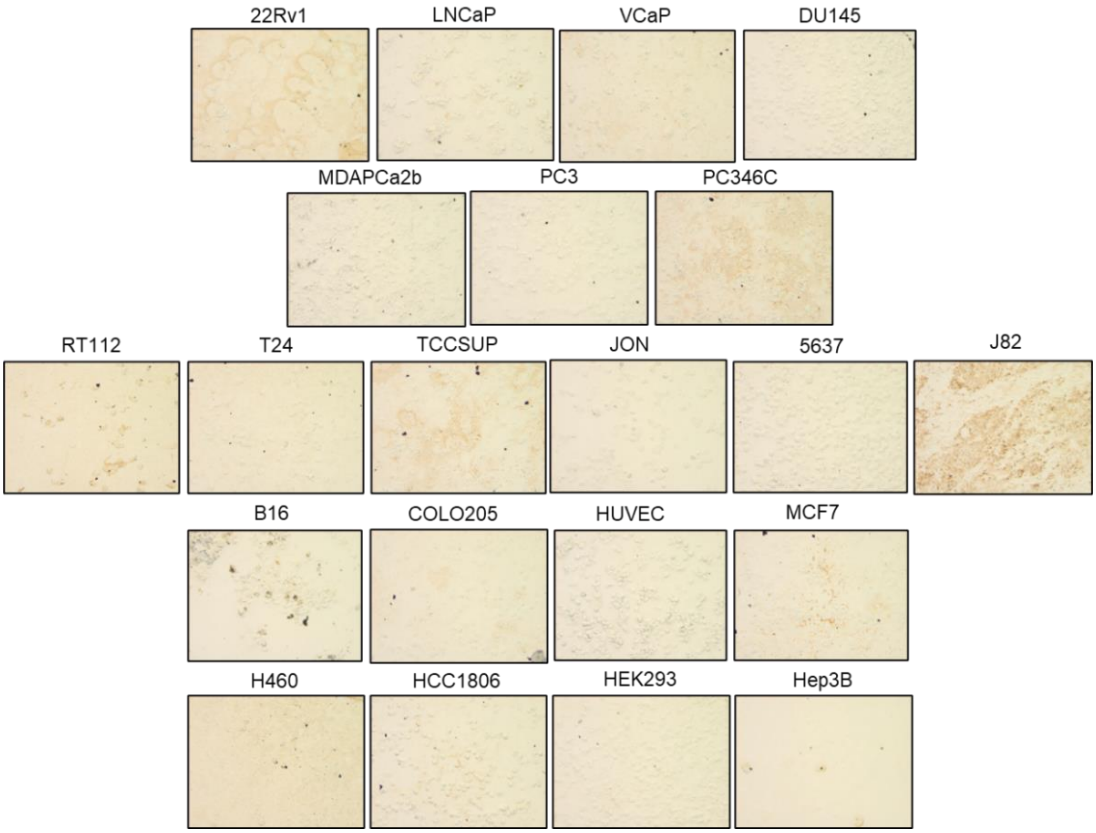


Figure 53 | nonBlCaNb1 Nb binding to different Ca cell lines.

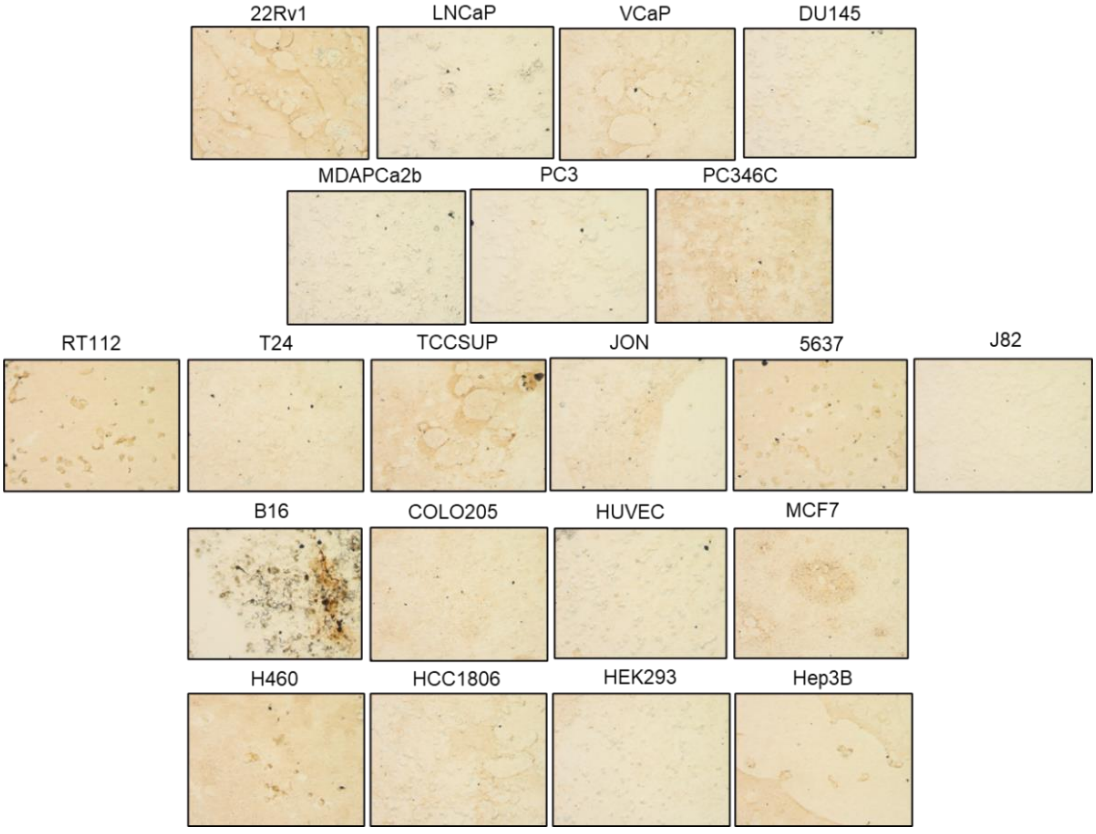


Figure 54 | nonBlCaNb2 Nb binding to different Ca cell lines.

7 FroCellCA BlCaNb35, BlCaNb36, BlCaNb40, BlCaNb41 at higher concentrations

Figure 55 – 58 | Images of DAB staining showing the affinity between BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 (10 x 10^{10} ø/mL) and twenty different cell lines. For immunocytochemical staining, we included six BlCa cell lines (T24, TCCSUP, RT112, J82, JON, 5637), seven PCa cell lines (22Rv1, DU145, LNCaP, MDAPCa2b, PC3, PC346C, VCaP) and other Ca cell lines (COLO205, H460, HCC1806, HEK293, Hep3B, HUVEC and MCF7). Images were taken at 20x magnification using the BX41TF Olympus microscope.

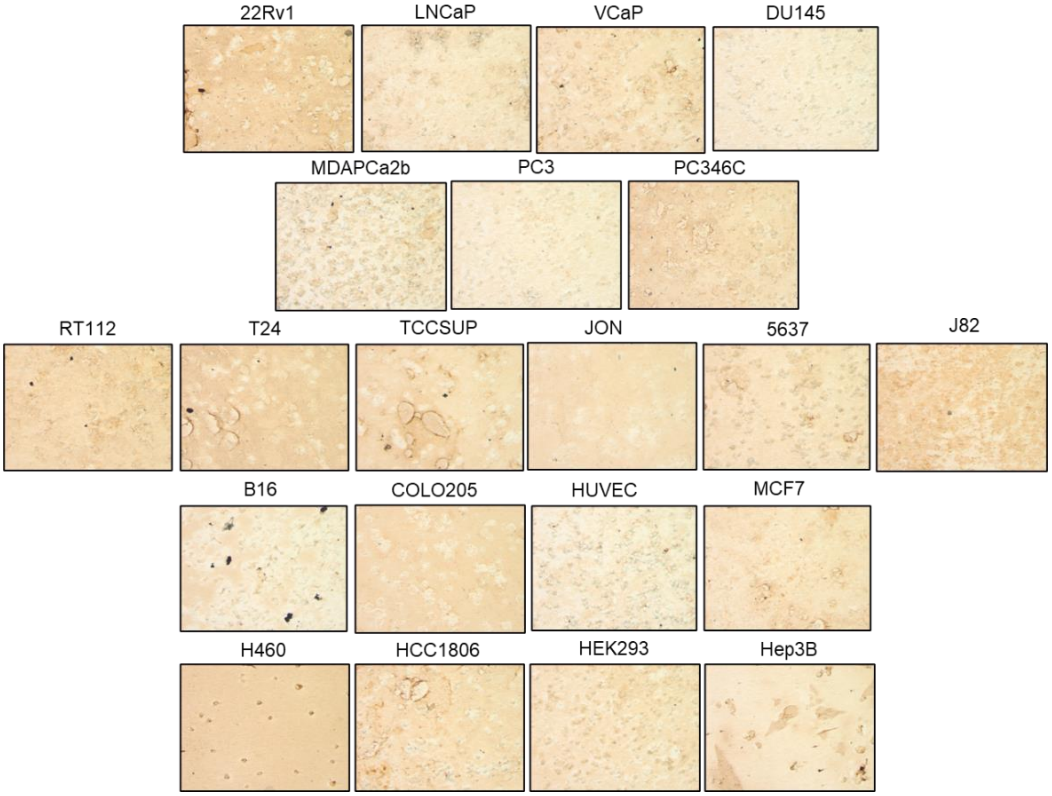


Figure 55 | BlCaNb35 Nb binding to different Ca cell lines.

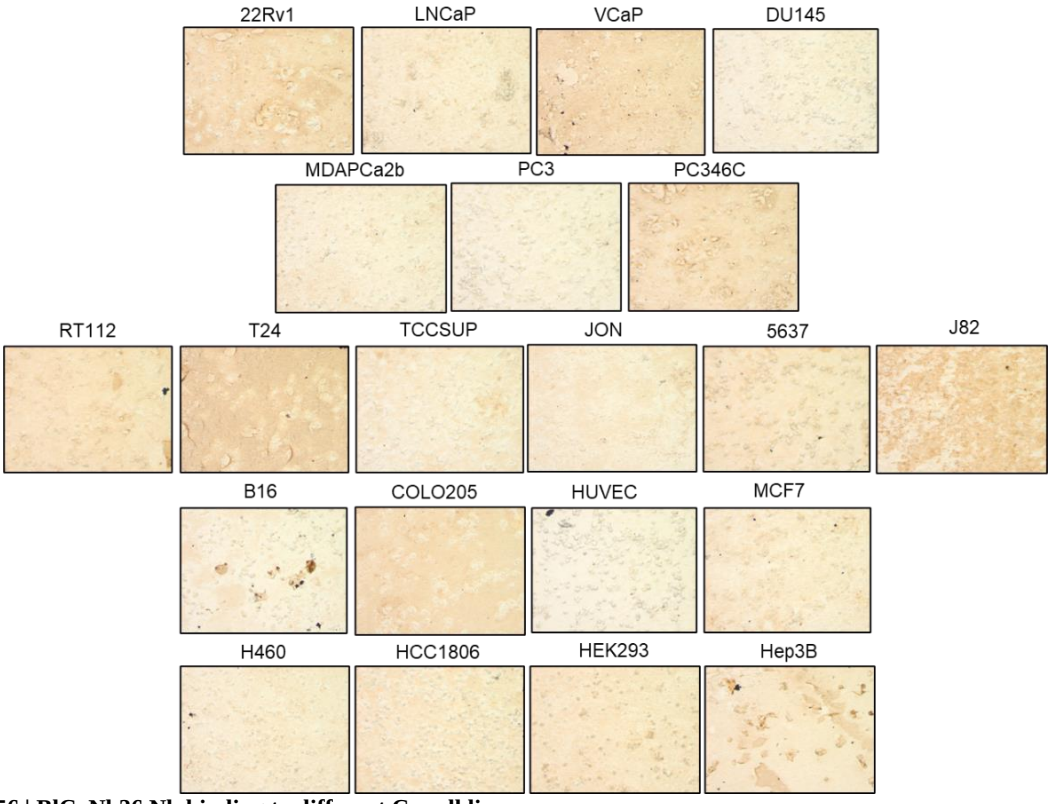


Figure 56 | BlCaNb36 Nb binding to different Ca cell lines.

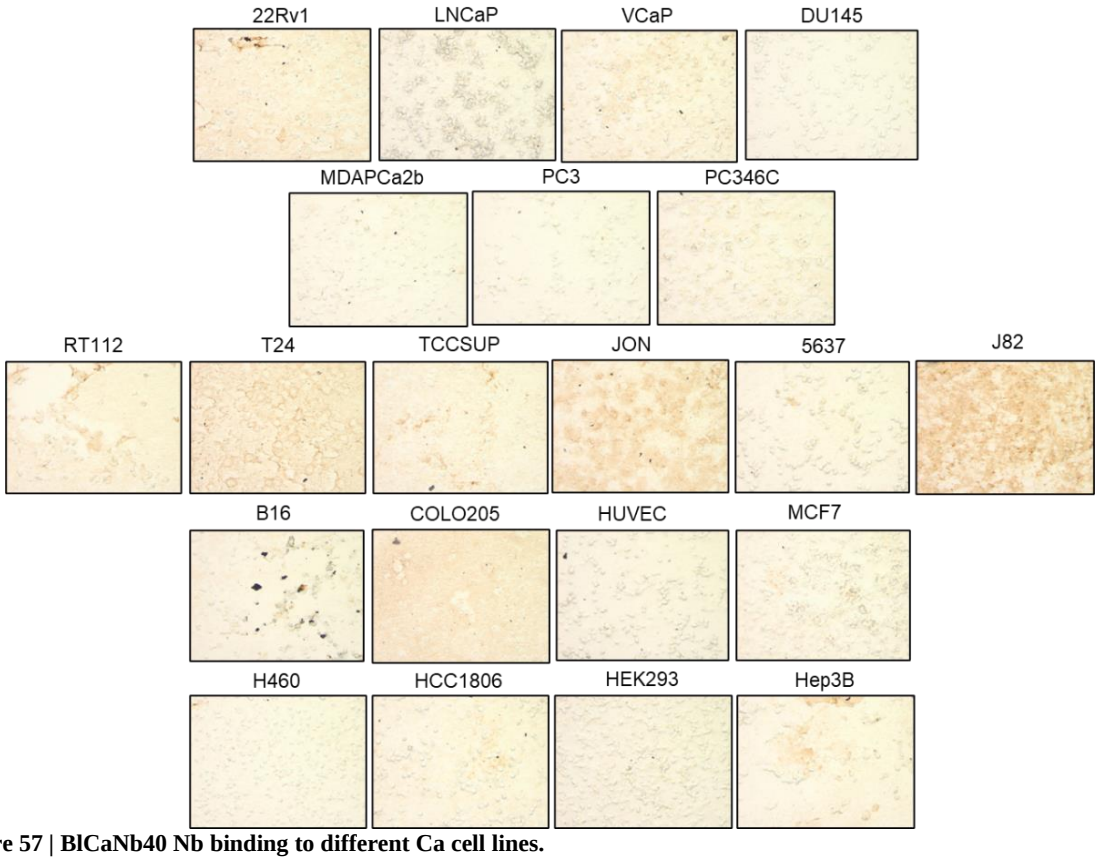


Figure 57 | BlCaNb40 Nb binding to different Ca cell lines.

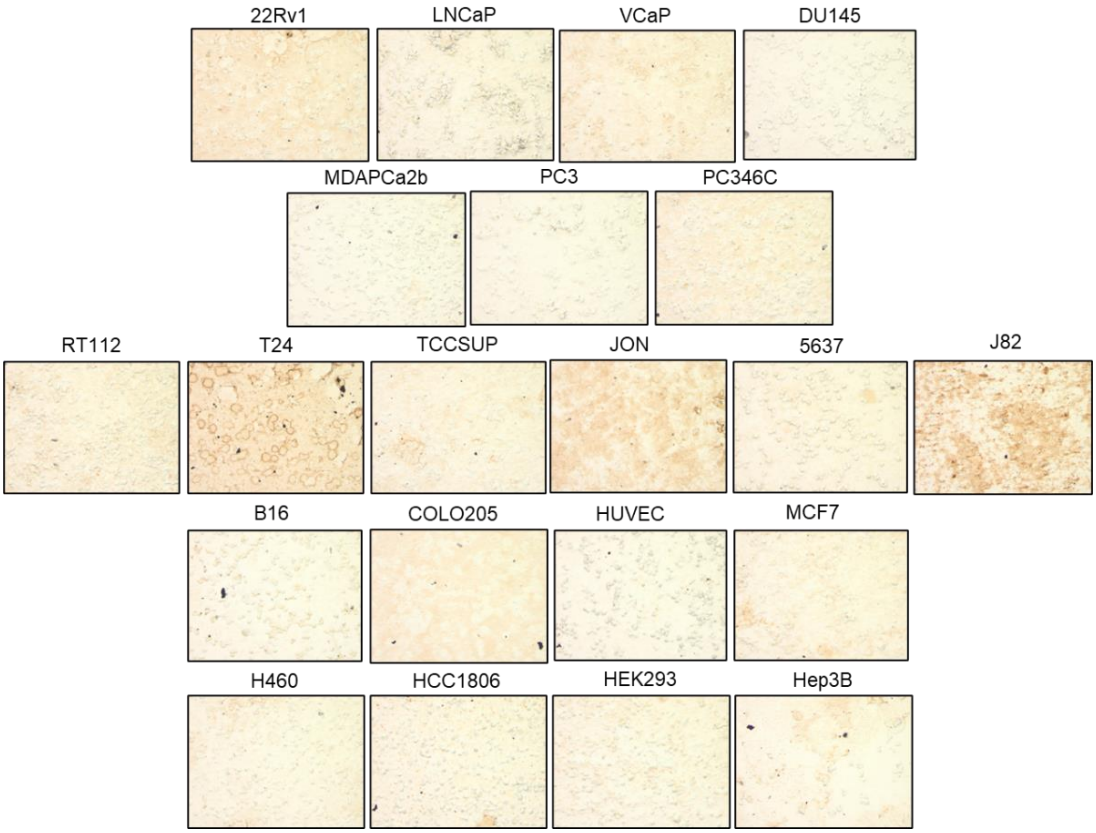
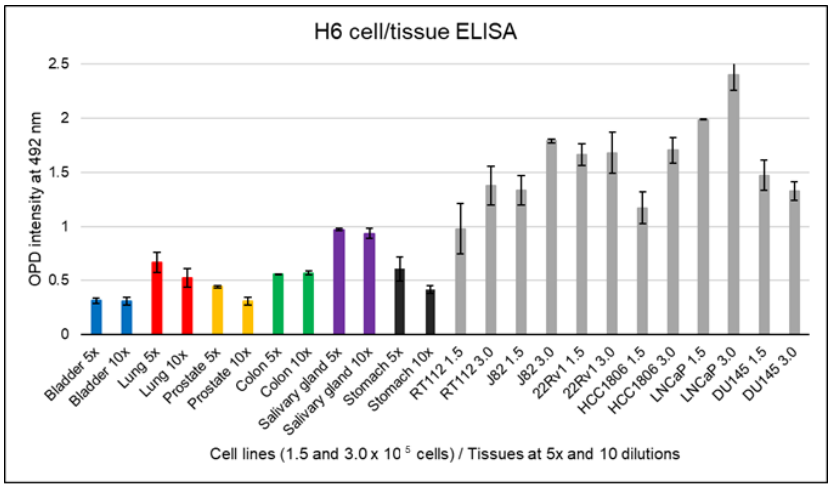
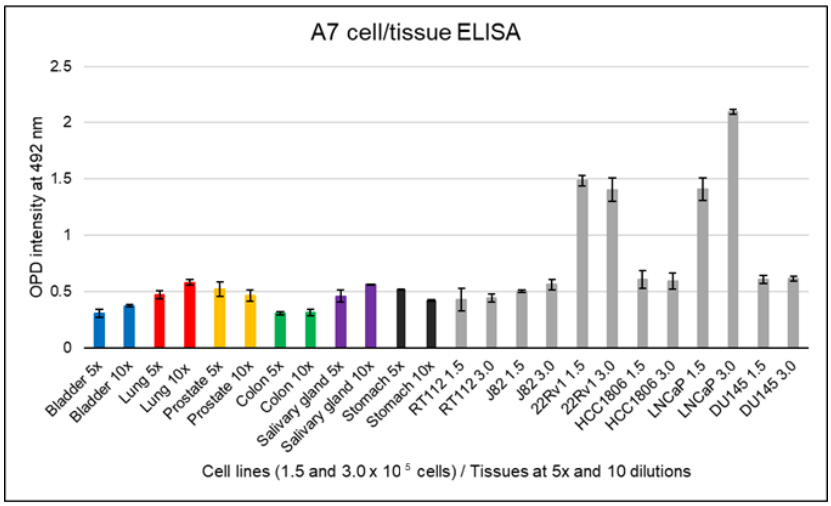


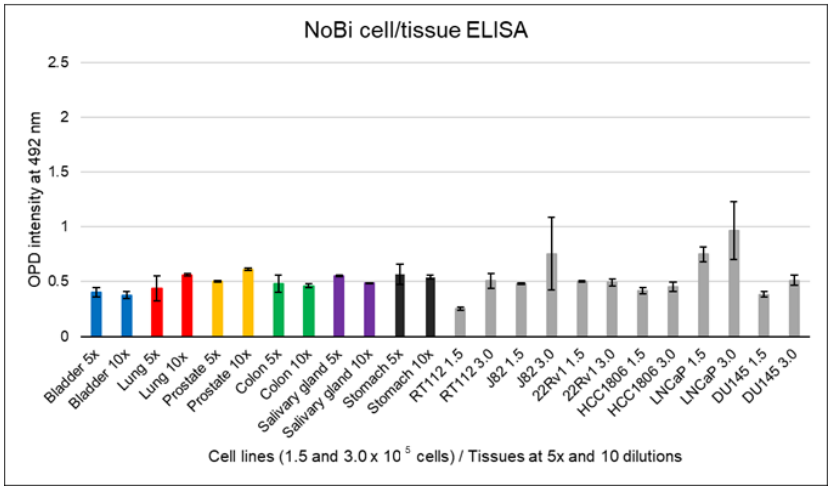
Figure 58 | BlCaNb41 Nb binding to different Ca cell lines.

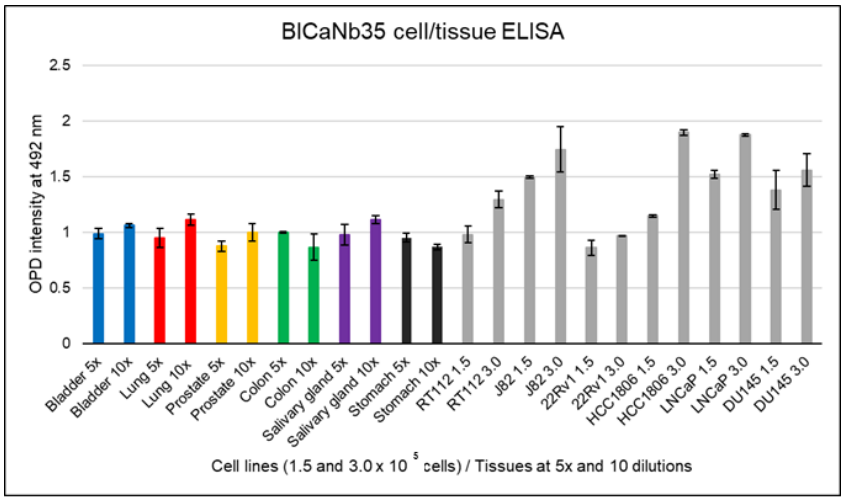
8 Cell/tissue ELCA plots

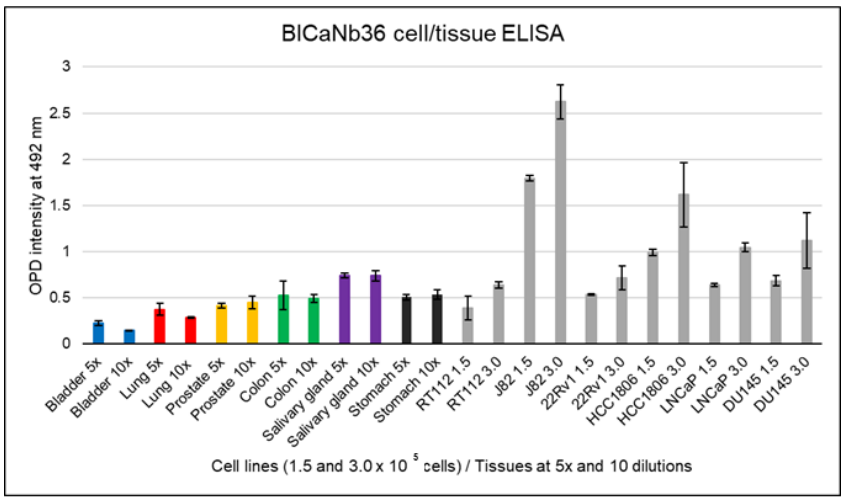
Figure 59 | Nb binding to cells/tissue using enzyme-linked cell assay (ELCA). From top to bottom, the binding activity of Nbs H6, A7, NoBi, BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 is shown against healthy bladder tissue (blue), lung tissue (red), prostate tissue (orange), colon tissue (green), salivary gland tissue (purple), stomach tissue (black), and several cell lines (grey). Binding levels were tested when added 5 x 10^{10} ø/mL phages to tissues at 5x and 10x dilutions from the original vial and to 1.5 x 10^{5} and 3.0 x 10^{5} cells. All samples were included in duplicates during ELCA, the averages are plotted and the error bar (standard deviation) is shown for each sample.

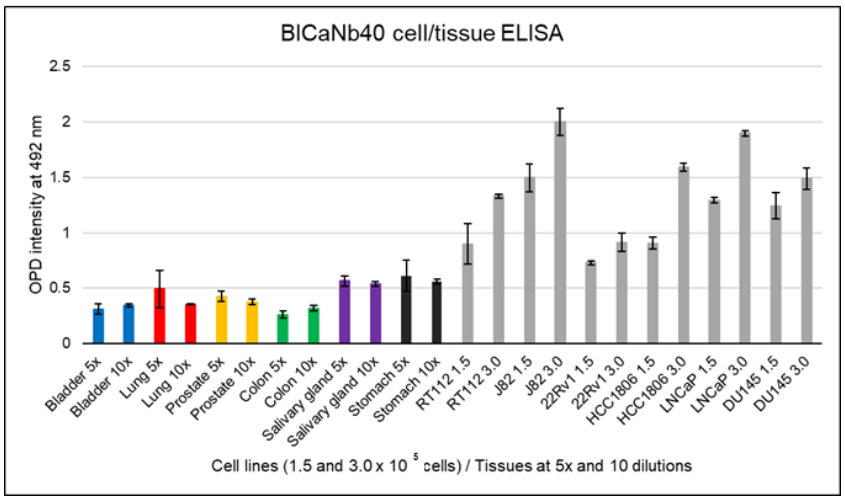


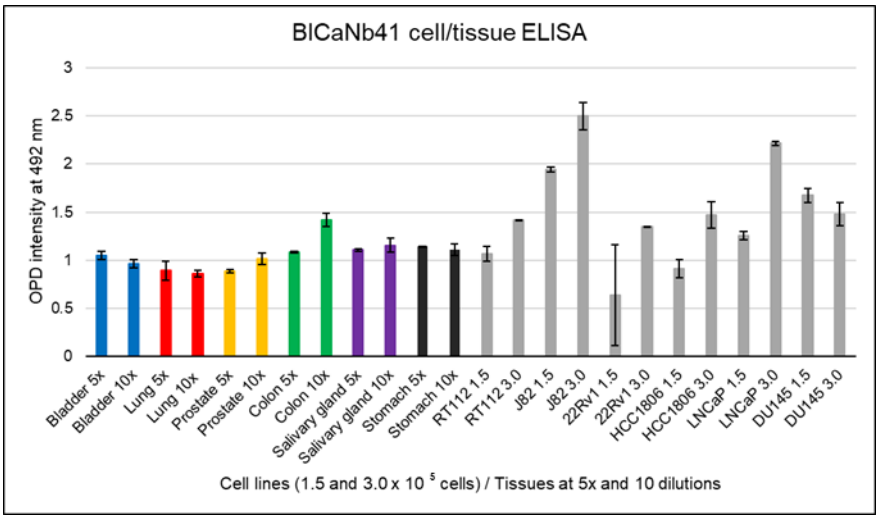





BICaNb35 cell/tissue ELISA
OPD intensity at 492 nm
2.5
2
1.5
1
0.5
0
Bladder 5x
Bladder 10x
Lung 5x
Lung 10x
Prostate 5x
Prostate 10x
Colon 5x
Colon 10x
Salivary gland 5x
Salivary gland 10x
Stomach 5x
Stomach 10x
RT112 1.5
RT112 3.0
J82 1.5
J82 3.0
22Rv1 1.5
22Rv1 3.0
HCC1806 1.5
HCC1806 3.0
LNCaP 1.5
LNCaP 3.0
DU145 1.5
DU145 3.0
Cell lines (1.5 and 3.0 x 10 5 cells) / Tissues at 5x and 10 dilutions


BICaNb36 cell/tissue ELISA
OPD intensity at 492 nm
3
2.5
2
1.5
1
0.5
0
Bladder 5x
Bladder 10x
Lung 5x
Lung 10x
Prostate 5x
Prostate 10x
Colon 5x
Colon 10x
Salivary gland 5x
Salivary gland 10x
Stomach 5x
Stomach 10x
RT112 1.5
RT112 3.0
J82 1.5
J82 3.0
22Rv1 1.5
22Rv1 3.0
HCC1806 1.5
HCC1806 3.0
LNCaP 1.5
LNCaP 3.0
DU145 1.5
DU145 3.0
Cell lines (1.5 and 3.0 x 10 5 cells) / Tissues at 5x and 10 dilutions


BICaNb40 cell/tissue ELISA
OPD intensity at 492 nm
2.5
2
1.5
1
0.5
0
Bladder 5x
Bladder 10x
Lung 5x
Lung 10x
Prostate 5x
Prostate 10x
Colon 5x
Colon 10x
Salivary gland 5x
Salivary gland 10x
Stomach 5x
Stomach 10x
RT112 1.5
RT112 3.0
J82 1.5
J82 3.0
22Rv1 1.5
22Rv1 3.0
HCC1806 1.5
HCC1806 3.0
LNCaP 1.5
LNCaP 3.0
DU145 1.5
DU145 3.0
Cell lines (1.5 and 3.0 x 10 5 cells) / Tissues at 5x and 10 dilutions


BlCaNb41 cell/tissue ELISA
OPD intensity at 492 nm
3
2.5
2
1.5
1
0.5
0
Bladder 5x
Bladder 10x
Lung 5x
Lung 10x
Prostate 5x
Prostate 10x
Colon 5x
Colon 10x
Salivary gland 5x
Salivary gland 10x
Stomach 5x
Stomach 10x
RT112 1.5
RT112 3.0
J82 1.5
J82 3.0
22Rv1 1.5
22Rv1 3.0
HCC1806 1.5
HCC1806 3.0
LNCaP 1.5
LNCaP 3.0
DU145 1.5
DU145 3.0
Cell lines (1.5 and 3.0 x 10 5 cells) / Tissues at 5x and 10 dilutions

9 Cell ELCA plots with B16 correction

Figure 60 | Nb binding to cells using enzyme-linked cell assay (ELCA). From top to bottom, the binding activity of Nbs BlCaNb35, BlCaNb36, BlCaNb40 and BlCaNb41 is shown against different cell lines. Binding levels were tested when added 5 x 10^{10} ø/mL phages to 3.0 x 10^{4} and 6.0 x 10^{4} cells. All samples were included in duplicates during ELCA, the averages are plotted and the error bar (standard deviation) is shown for each sample. The plots were corrected for the averaged signal against B16 3 and B16 6 for each BlCaNb separately.