Development and applications of novel approaches for monitoring antibody generation to prostate cancer–related antigens

Darragh Lemass, B. Sc. (Hons.)

This thesis is submitted to Dublin City University for the degree of Ph.D.

Based on research carried out at

School of Biotechnology,

Dublin City University,

Dublin 9, Ireland.

Supervisors: Professor Richard O'Kennedy

Dr. Gregor Kijanka

July 2017

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Ph.D. is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ ID No.: _____ Date: _____

Acknowledgements

The work presented in this thesis would not have been achievable without the support of many people. First I would like to thank Professor Richard O'Kennedy for his motivation, guidance and patience, which were ever-present over the course of this research. I also wish to thank Dr. Gregor Kijanka, for his enthusiasm, encouragement and friendship. I am indebted to both of you, for the meticulous supervision, expertise and time that you have contributed to this research.

I also wish to acknowledge Dr. Hui Ma of the Applied Biochemistry Group as well as Ray Moran and Dr. Mary O'Connell of the Computational & Molecular Evolutionary Biology Group, for their valued contributions to the work presented here.

To my colleagues in the Applied Biochemistry Group and in DCU, I feel privileged to have worked alongside so many great friends. Be it the banter in the lab, escaping the bench for a coffee or drinks in town, you made the bad days seem few and far between. I also wish to thank you for the technical advice, which was always offered freely and with a smile.

Thank you to my friends outside of DCU. Cian, Daragh, Séan, Adrian, Mark and so many others, thank you for the laughs, the mad nights and the games of AstroTurf and FIFA tournaments. All the very welcome distractions have kept me sane over the years.

Finally, I wish to thank my parents Peter and Helen, for whom I dedicate this thesis, my siblings Cian, Kate and Claire, and my girlfriend and best friend Sinéad. Your unwavering love and support got me through the tough times and kept things in perspective. Without you this work would not have been completed.

Table of Contents

List of figures

Chapter 1 - Introduction

Chapter 2 - Materials and Methods

Chapter 3 - Generation of a recombinant antibody library against Prostate Cancer cells

Chapter 4 - High density protein arrays for monitoring processes in recombinant antibody generation

Chapter 5 - Characterisation of scFv clones for

protein array-identified antigens

List of Tables

List of abbreviations

AA	Amino acid
Ab	Antibody
ADP	Adenosine diphosphate
AP	Alkaline phosphatase
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BLAST	Basic local alignment tool
BRU	Bio-resource unit
BSA	Bovine serum albumin
CAT	Computed tomography of the body
cDNA	Coding DNA
CDR	Complementarity determining region
C_H	Constant heavy chain
CM	Carboxymethylated
D	Diverse segment
DAVID	Database for annotation, visualization and integrated discovery
DHT	Dihydrotestosterone
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline (magnesium and calcium-free)
DRE	Digital rectal examination

EDC	Ethyl(dimethylaminopropyl)-carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-Linked Immunosorbent Assay
ENST	Ensembl transcript
ERG	ETS-related gene
Fab	Antigen-binding fragment
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
Fc	Fragment crystallisable
FC	Flow cell
GuHCl	Guanidine hydrochloride
HA	Haemagglutinin
HAT	Hypoxanthine, aminopterin and thymidine
Her2/neu	Receptor tyrosine-protein kinase erbB-2
HGPRT	Hypoxanthine-guanine phosphoribosyltransferase
His_6	Histidine x 6
HRP	Horse radish peroxidase
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgG1	Immunoglobulin G subclass 1
IgG2	Immunoglobulin G subclass 2

IgG4	Immunoglobulin G subclass 4
IgM	Immunoglobulin M
IgY	Immunoglobulin Y
IHC	Immunohistochemical
IMAC	Immobilised metal affinity chromatography
ITPG	Isopropyl β-D-1-thiogalactopyranoside
J	Joining segment
LacZ	*LacZ* β-galactosidase
LAST	Laboratory animal science and training
LB	Luria Bertani broth
mAb	Monoclonal antibody
MM	Milk marvel
MS	Mass spectrometry
MW	Molecular weight
NCBI	National Centre for Biotechnology Information
NHS	N-hydroxysulphosuccinimide
NTA	Nitrilotriacetic acid
OD	Optical density
P3P2F10	Pan 3 output, Plate 2, Well F10
pAb	Polyclonal antibody
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with tween
PCa	Prostate cancer
PCR	Polymerase chain reaction

PSA	Prostate specific antigen
PEG	Polyethylene glycol
PSMA7	Proteasome subunit alpha type-7
PVDF	Polyvinylidene fluoride
RANKL	Receptor activator of nuclear factor kappa-B ligand
RARG	Retinoic acid receptor gamma
RASdmi	Ras, dexamethasone-induced 1
RIPA	Radio-immunoprecipitation assay
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
TAE	Tris acetate EDTA
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline with tween
TEMED	N,N,N',N'-Tetramethylethylenediamine
TMB	3,3’,5,5’-Tetramethylbenzidine
TNM	Tumour node metastasis
TRUS	Trans rectal ultrasound
SB	Super broth
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SETD2	SET domain containing 2
SOC	Super optimal broth with catabolite repression
SOE	Splice by overlap extension
SS	Stuffer fragment

STR	Short tandem repeats
scAb	Single-chain antibody fragment
scFv	Single-chain variable fragment
SPR	Surface plasmon resonance
V	Variable segment
ZNF358	Zinc finger protein 358

List of units

%	Percentage
X^2	Chi^2
^oC	Degrees Celsius
bp	Base pair
Da	Daltons
g	Grams
µ	Micro-
m	metre
c	Centi-
k	Kilo-
L	Litre
m	Milli-
min	Minutes
M	Molar
n	Nano-
OD	Optical density
pH	Negative logarithm of the hydrogen ion concentration
RU	Response units
Rpm	Revolutions per minute
v/v	Volume per unit volume
sec	Seconds
w/v	Weight per unit volume
x g	Relative centrifugal force

Publications and presentations

Publications

Lemass, D., Hearty, S. and O'Kennedy, R. (2014) Surface Plasmon Resonance-Based Immunoassays, in: Koistinen, H. and Stenman, U.-H. (eds.) *Novel Approaches in Immunoassays*. London: Future Science Group, pp. 20–40.

Lemass, D., O'Kennedy, R. and Kijanka, G. S. (2016) *Referencing cross-reactivity of detection antibodies for protein array experiments*, F1000 Research, 5, pp. 73. (Manuscript in open review).

Moran, K. L. M., **Lemass, D.** and O'Kennedy, R. (2017) *Surface Plasmon Resonance-based Immunoassays: Approaches, Performances and Applications*, in: Handbook of Immunoassay Technologies: Approaches, Performances, and Applications. Elsevier Ltd. (Manuscript accepted for publication).

Lemass, D., Moran, R. J., Ma, H., O'Connell, M. J., O'Kennedy, R. and Kijanka, G. S. (2017) Protein array-based monitoring of recombinant antibody generation to prostate cancer-related antigens. (Manuscript in preparation for submission).

Presentations

Lemass, D., O'Reilly, J. A. and O'Kennedy, R. *'In Vitro Technologies'* PROSENSE Workshop 2nd-3rd July 2013, Dublin City University, Dublin, Ireland. (Poster presentation).

Lemass, D., O'Kennedy, R. and Kijanka, G. S. *Rapid, Single-step Characterisation of Monoclonal scFv Pools from an Avian Phage Library Raised and Enriched Against Prostate Cancer*. The 7th Annual PEGS Europe Summit, 2nd-6th November 2015, Epic Sana Lisboa Hotel, Lisbon, Portugal. (Poster Presentation).

Lemass, D., O'Kennedy, R. and Kijanka, G. S. *Generation of Recombinant Antibodies to Prostate Cancer-Associated Antigens*.The 8th Annual School of Biotechnology Research Day, 29th January 2016, Dublin City University, Dublin, Ireland. (Oral Presentation).

Abstract

Darragh Lemass

Development and applications of novel approaches for monitoring antibody generation to prostate cancer–related antigens

Antibody–based approaches have provided vast improvements in the detection and treatment of many cancers. However, progress is currently limited in prostate cancer due to a lack of clinically-relevant biomarkers. Novel biomarker discovery may be achieved with antibody technology, which utilises cancer cells to pan antibody libraries/sources, to identify antibodies against non-predetermined antigens. The major limitation currently encountered is the difficulty in establishing the identities of such antigens, due to the need for extensive post-selection antibody characterisation.

The work embodied in this thesis utilised protein array technology to monitor the generation of a large recombinant antibody library directed towards non-predetermined antigens present in a prostate cancer cell line. Protein arrays were used to profile the immune response, mounted by the avian host following cancer cell immunisation. Investigations of the antigen profile revealed a trend in the immune response towards orthologs that were conserved between human and chicken. Similar protein array-profiling was utilised to characterise the resulting avian scFv-phage library generated, as well as 384 scFv isolated following enrichment of the library towards

prostate cancer cells. Evaluation of identified scFv-antigens allowed for the selection of an antigen subset to be screened against the scFv library. This screening resulted in the isolation of the first recombinant antibody to Zinc Finger Protein 358, an antigen currently uncharacterised in prostate cancer. This antibody was isolated and characterised.

In conclusion, the incorporation of protein array technology into recombinant antibody-based biomarker discovery enabled antigen characterisation at an earlier stage and on a larger scale than possible with traditional methods, thus circumventing the bottleneck associated with the current strategies in biomarker discovery using recombinant antibody technology. The high-throughput approach developed in this study sheds light on our understanding of immunisation processes and substantially enhances the potential of recombinant antibody technology for cancer biomarker discovery and antibody development.

Chapter 1 - Introduction

1.0 Section overview

Prostate cancer (PCa) is the most diagnosed form of cancer in men. However, the current methods of detecting and treating prostate PCa are ineffective. Antibodies are excellent tools for diagnosing and treating cancer, yet, established generation strategies rely on the identification and extensive characterisation of targets prior to development. The difficulty in identifying cancer-specific biomarkers acts as a bottleneck to improving diagnostic assays and therapeutic approaches. This chapter introduces the methods of generating monoclonal, polyclonal and recombinant antibodies, as well as the present landscape of PCa detection and treatment. In addition, the current approaches to biomarker discovery including proteomics, phage display and classical methods are discussed. The chapter closes with an outline of the goals of this research, which aim to improve the current strategies utilised in generating antibodies to non-predetermined antigens, to aid biomarker discovery. This is to be achieved by augmenting the current phage display strategy with protein array technology.

1.1 Antibodies

1.1.1 The immune system

The immune system is the body's mechanism of defence against pathogens such as bacteria, fungi and viruses. The immune system's response to a pathogen can be divided into two categories: innate and acquired.

Innate immunity acts as the body's primary response to pathogenic invaders. The mechanisms of action have evolved to counteract all invaders, rather than specific pathogens. The innate immune system starts from the skin, mucosal membranes and the anti-microbial saliva and tear ducts - acting as an initial barrier to infection. If a pathogen bypasses these barriers, the innate immune system is activated. Typically, a macrophage will recognise the invader by its 'pathogen-associated' molecular pattern and will release a range of mediators, such cytokines and chemokines, to initiate an inflammatory response. This directs further effector cells, primarily neutrophils and macrophages, to the site of infection before an adaptive immune response is activated. The innate response is vital for containing the early stages of infection, as it takes 4-7 days to mount an adaptive immune response (Janeway *et al.*, 2001).

While innate immunity plays an important role in the initiation of an immune response, completion of the immune response often falls under the remit of the adaptive immune system (when additional immune responses are necessary to successfully combat the infection). An important occurrence during the innate response is the transport of phagocytosed pathogens to the lymph nodes by activated dendritic cells. In the lymph nodes, these dendritic cells, also known as antigen-presenting cells, display fragments

of the ingested pathogen on their surface with major histocompatibility complex proteins. The lymph nodes contain naïve lymphocytes. Lymphocytes are B-cells or T-cells and mediate humoral or cell-mediated immunity respectively. Unlike innate immune cells, lymphocytes target one molecule associated with a pathogen, known as an antigen. The immune system has the capacity to recognise 10^8 unique antigens *in vivo* (Conroy *et al.*, 2014). While cell-mediated immunity mostly uses T-cell receptors for antigen recognition, antibodies are the recognition mechanism of the humoral response. Antibodies bind to their specific antigen and mark the antigen-displaying pathogen for destruction. Clearance of antigen-coated pathogens can be achieved with phagocytic cells or the complement system.

To produce an effective level of antibodies to combat an infection, a naïve lymphocyte must first be activated by encountering its specific antigen. The binding of an antigen to a B-cell receptor, in the presence of helper T-cells, induces maturation and proliferation. B-cells can mature into antibody-secreting plasma cells or memory cells. Helper T-cells are also recruited to the site of infection, where they secrete stimulatory cytokines and direct cytotoxic T-cells, plasma cells and phagocytes to the pathogen. Cytotoxic T-cells present receptors on their surface capable of recognising a specific antigen, and will eliminate any infected cells displaying that antigen on their surface. Following destruction of a pathogen, memory B-cells and T-cells remain in the body for rapid response in the case of re-infection – granting the host immunity. Memory cells undergo rapid clonal expansion should they encounter the antigen again, producing a specific immune response against the pathogen and granting immunity to the host. A simplified schematic of the innate-adaptive immunity link is presented in Figure 1.1.

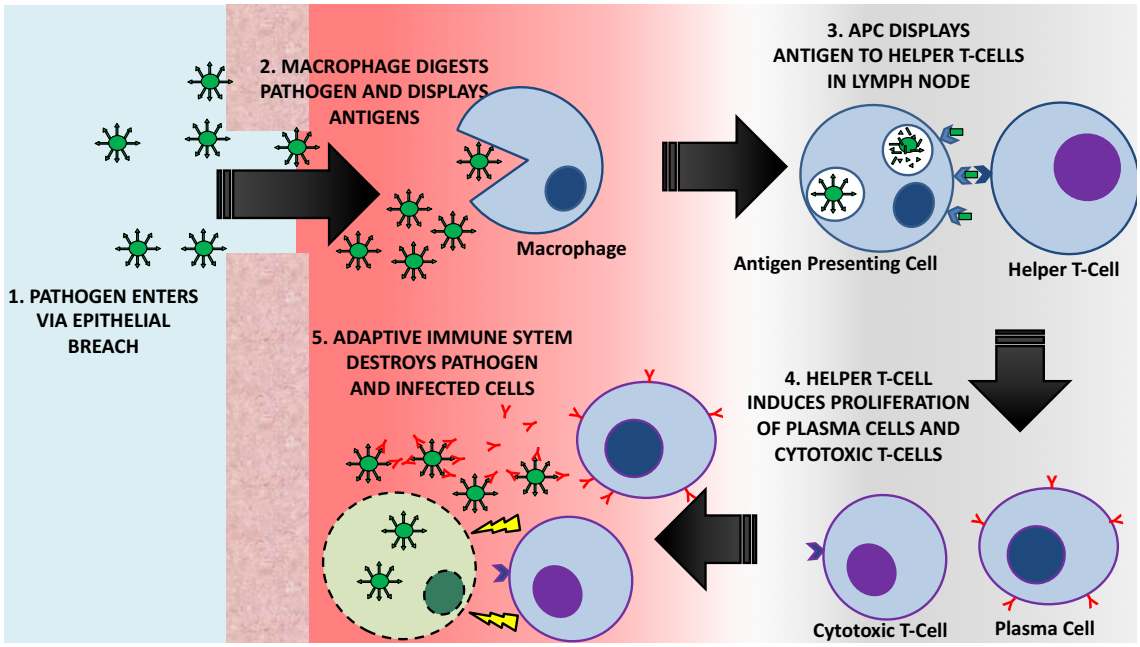

Figure 1.1 - Summary of the immune response to an infection. 1. The infection process starts when a pathogen breaches a barrier of the host. 2. Phagocytic cells of the innate immune system recognise and endocytose the pathogen. 3. Phagocytic cells migrate to the lymph nodes as antigen presenting cells. Antigen presenting cells activate helper T-cells to initiate the adaptive immune response. 4. Helper T-cells induce the maturation and proliferation of plasma cells and cytotoxic T-cells specific to the antigen. 5. Mature plasma and cytotoxic T-cells migrate to the sites of infection. Cytotoxic T-cells destroy pathogen-harbouring host cells. Plasma cells secrete antibodies to immobilise pathogens and mark them for destruction.

1.1.2 Antibody structure

Antibodies or immunoglobulins (Ig) are glycoproteins produced by the body in response to the presence of antigens and act as the key recognition elements of the adaptive immune system. They are found either solubilized in the blood or tissue matrices, or on the surface of B-cells, as B-cell receptors. Antibody-binding will typically neutralize a foreign antigen and mark it for destruction by the immune system.

IgG is representative of the basic structure of an antibody. IgG comprises four polypeptide chains – two identical heavy chains (~440 amino acids (AA) in length) and two identical light chains (~220 AA in length) held together with disulphide bonds (see Figure 1.2A). Each chain consists of 110 AA segments that are also held together by disulphide bonds (Woof and Burton, 2004). The heavy chains are made up of three constant domains (C_{H1}, C_{H2} and C_{H3}) and one variable domain (V_H). Both heavy chains are covalently linked by a pair of disulphide bonds located at a flexible amino acid (AA) hinge between the C_{H1} and C_{H2} regions. Single disulphide bonds link the light chains to the heavy chains above the hinge region (Barbas III *et al.*, 2001). The light chains are made up of one variable (V_L) and one constant domain (C_L). These chains combine with the V_H and C_{H1} segments to form the arms of the Y-shaped immunoglobulin, referred to as the antigen-binding fragments (Fab). The C_{H2} and C_{H3} segments make up the immunoglobulin's stem, known as a crystallisable fragment (Fc). The Fc region of the antibody determines its isotype and effector functions. Such functions include complement system activation and interaction with the Fc receptors expressed by phagocytic cells (Schroeder *et al.*, 2010). The Fab domains of an antibody are responsible for recognizing an antigen. Both the heavy and light variable domains consist of domain-pairing framework regions, allowing both chains to interact to form the fragment variable (Fv) region responsible for antigen binding. Three hyper variable loops are present in both the heavy and light chain. These hypervariable loops are also known as complementarity determining regions and are responsible for the immunoglobulin physically binding to its cognate antigen. The region of the antigen bound by the immunoglobulin is called the epitope. Complementarity determining regions vary significantly and are responsible for the adaptive immunity's ability to recognize a huge

number of pathogens (Barbas III *et al.*, 2001). Immunoglobulin gamma, or IgY, is the IgG-equivalent in birds. It is thought to be the ancestor of both IgG and IgE isoforms (Warr *et al.*, 1995). The major difference between IgY and IgG is the presence of an additional domain on the heavy chain, making IgY ~180 kDa in size. This domain has been proposed to be an evolutionary precursor to the hinge region of the IgG, meaning IgY is comparatively less flexible. Both IgG and IgY are depicted in Figure 1.2.

In humans, there are five isoforms of antibodies - IgG, IgA, IgM, IgE and IgD. Each isoform has a different heavy chain structure and all play a different role in immune function (Schroeder *et al.*, 2010). Despite the variation in heavy chain structure, all isoforms are made up of the same basic Y-shaped structure. In species other than humans, immunoglobulins can also be found naturally in alternative formats. Examples include the IgY format in birds, the nanobody format in camelids and the Ig New Antigen Receptor format present in sharks (Dias da Silva and Tambourgi, 2010; Krah *et al.*, 2016; Steeland *et al.*, 2016). The advent of recombinant antibody technology (discussed later) has introduced additional isoforms such as the single chain antibody variable region fragment (scFv) and the single chain antibody fragment (scAb) (see Figure 1.2C and 1.2D).

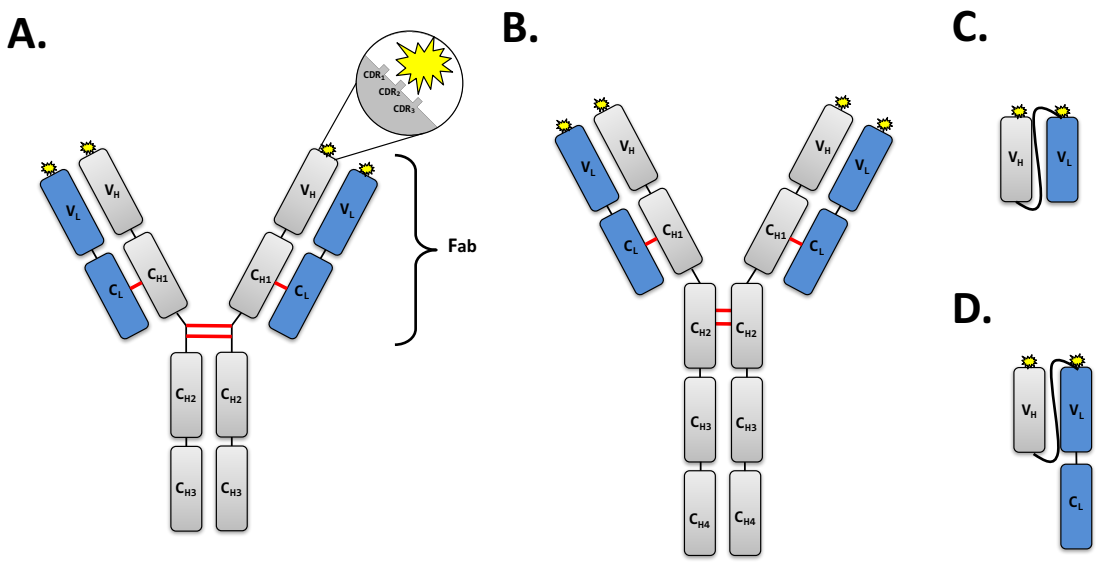

Figure 1.2 – Different immunoglobulin formats. Figure adapted from Lemass *et al.* (2014). Image A depicts an IgG molecule with a CDR region highlighted. Image B depicts an IgY molecule. Image C depicts a scFv molecule. Image D depicts a scAb molecule.

A host's range of antibody specificity is estimated to be sufficient to cover the entire antigen repertoire. Diversity in specificity is conferred by genetics. In mammals, such as humans, immunoglobulin diversity is achieved via gene rearrangement where multiple functional Variable (V), Diverse (D) and Joining (J) segments are available to form V_LJ and $V_HJ(D)$ cDNA for light and heavy variable chains. This diversity is further augmented by a phenomenon known as somatic hypermutation (Dias da Silva and Tambourgi, 2010). Somatic hypermutation occurs in B-cells, in which point mutations are inserted into the already assembled Ig variable gene segments at rate ~10^6 times higher than naturally-occurring mutations. In avian species, such as chickens, the mechanism granting diversity to immunoglobulins is less complex. Upstream of single V genes are pseudogene blocks, for both the heavy and light chains. A random selection of the pseudogenes are transplanted into the variable genes to achieve diversity (Ratcliffe,

2008; Wu *et al.*, 2012). A comparison of both approaches to achieve diversity is depicted in Figure 1.3.

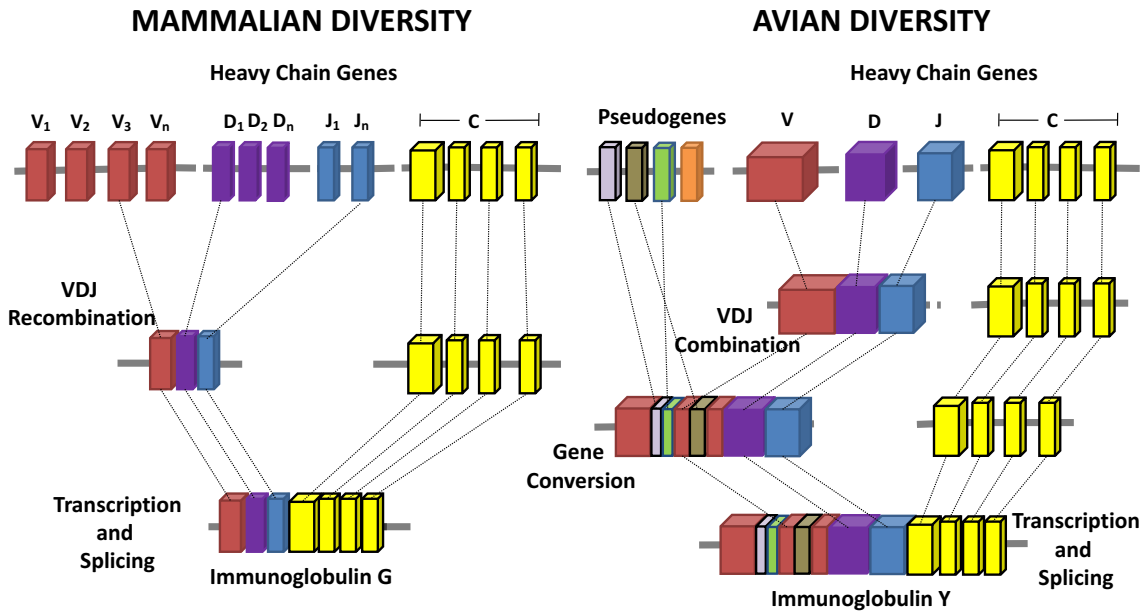

Figure 1.3 – Comparison of genetic immunoglobulin diversity mechanisms employed in mammalian and avian immune systems. Both examples illustrate diversity in the heavy chains, though the same approach is also used to achieve light chain diversity in both systems. Mammalian systems (left half of image) achieve diversity through somatic recombination, where multiple VDJ segments (red, purple and blue genes) are recombined in immature B-cells for diversity, and linked by constant genes (yellow) to determine the immunoglobulin isotype. Avian systems (right half of image) have single variable genes and instead achieve diversity through gene conversion. In gene conversion, upstream pseudogene segments are translocated into the variable gene before transcription. The VDJ segments are then linked to constant genes to determine the immunoglobulin's isotype.

1.1.3 Methods of antibody generation

Antibodies can be either polyclonal or monoclonal. As the name suggests, polyclonal antiserum contains antibodies from more than one source whereas monoclonal antibodies are all genetically identical. The following section describes the currently used methods of generating antibodies against a target of interest. A diagram

summarising polyclonal, monoclonal and recombinant antibody generation is presented in Figure 1.5.

1.1.3.1 Polyclonal antibody preparations

Polyclonal antibodies are used extensively for research purposes and have been successfully employed in biosensors. The labour required to produce polyclonal antibodies is far less than that for monoclonal antibodies, yet polyclonal antibodies have significant drawbacks. Polyclonal antibodies are derived from multiple plasma cells and are therefore not genetically identical. Thus, a polyclonal antibody preparation may bind multiple epitopes on the same antigen. To generate a polyclonal antibody against an identified target-of-interest, a host undergoes an immunisation regime against that target, until an immune response is mounted. The polyclonal antibodies are then purified, primarily from the host's serum or egg yolk (if an avian host is used). When a new host is employed to produce another batch, the differences in the plasma cell repertoire between individuals can result in batch-to-batch variation. Hence polyclonal antibodies are not ideal for clinical purposes.

1.1.3.2 Hybridoma technology

A major step in the ability to generate antibodies was taken by Köhler and Milstein (1975), when they pioneered hybridoma technology. Hybridomas are single cell lines, capable of producing antibodies. They are made by fusing a lymphocyte producing the antibody of interest with a myeloma cell. This is achieved using polyethylene glycol, which promotes cell fusion and exchange of nuclei. Following fusion of lymphoctyes with myeloma cells, the cells are cultured in HAT (Hypoxanthine, Aminopterin and

Thymidine)-containing medium. HAT-containing medium selects for cells which contain hypoxanthine-guanine phosphoribosyltransferase (HGPRT), which is present in the lymphoctyes but not in the myeloma cells. The aminopterin in the medium inhibits DNA *de novo* synthesis in HGPRT-lacking cells. Thus, non-fused myeloma cells will not survive in HAT-media and die. Lymphocytes fused with myeloma cells (i.e. hybridomas) can continue to grow. Hybridomas are then isolated and characterised (Little *et al.*, 2000). Antibodies produced by hybridomas are genetically identical and hence bind the same epitope.

Immunisation of an animal, typically a mouse, with an antigen of interest and isolation of hybridomas is now an established method to generate valuable antibodies. A significant limitation of the technology is the screening process, downstream of hybridoma generation and isolation of individual cells. Typically, monoclonal hybridomas are grown in microplates for characterisation. The antibodies expressed by each hybridoma are generally screened by ELISA (Cervino *et al.*, 2008). The introduction of high-throughput robotic screening platforms, such as the Biacore 4000, may be used to improve the output from screening in terms of the quality of the antibodies selected (e.g. their affinity, stability) (Conroy *et al.*, 2012). Further improvements to the technology are continuous; for example, a recent advance has augmented the technology with microfluidics, increasing the screening capacity to aid discovery (Fitzgerald *et al.,* 2015b).

1.1.3.3.1 Recombinant antibody technology

The advent of recombinant antibody technology has greatly improved the capacity for generating antigen-binding fragments. Recombinant antibodies have multiple formats. Common formats are the single chain Fv (scFv) and single chain antibody fragment (scAb) (see Figure 1.2). All recombinant formats retain the variable region responsible for antigen-binding. The advantages offered by recombinant antibodies include ease of production, screening and the ability to engineer screened antibodies for the incorporation of tags and improvement of affinity. Recombinant antibodies have already been widely used in biosensors (Lemass *et al.,* 2014) and as biotherapeutics (Reichert, 2015). Another advantage offered by recombinant antibody technology is the ability to link two unique antibodies to make bifunctional antibodies – capable of binding to two separate analytes. This technology offers improved flexibility to the design of antibodies intended for diagnostic or therapeutic use (Byrne *et al.*, 2013).

Like hybridoma technology, recombinant antibody technology offers the capacity to utilise the immune response of a host against an antigen. It differs, however, in how individual antibody-producing clones are isolated. There are a number of approaches to achieve this, but most follow the same general strategy (Hoogenboom, 2005):

1 The generation of genotypic diversity

2 Linking the genotype to the phenotype

3 The application of an antigen-selective pressure to the phenotypic pool

4 Amplification of selected genotypes

For immune library generation, the RNA from the spleenocytes of a host (immunized or naïve) is isolated and converted into cDNA. The variable heavy and light gene repertoire

of the host is contained in this cDNA pool and holds the genetic information required to produce antibodies to the antigens to which the host was exposed. The host, however, will have generated an immune response to a vast number of antigens over its lifetime, which means that the mRNA pool will code antibodies against both the antigen-of-interest and many others. The challenge to antibody generation lies in isolating the genes of antibodies specific to the antigen of choice from this pool. To achieve this, display technology is employed. This links the genotype of an antibody to its phenotype, that can then be retrieved and amplified when a selective pressure is applied (Barbas III *et al.*, 2001). There are a number of ways to link an antibody's genotype to its phenotype. Validated display approaches include ribosome display, yeast display and phage display. Of these methods, phage display is the most commonly used due to its low cost, flexibility and overall robustness. (Hoogenboom, 2005).

1.1.3.3.2 Phage display

Phage display, first described in 1985 (Smith, 1985), uses bacteriophage to provide the phenotype-genotype antibody link in a library. Filamentous phage are engineered to display a scFv (or another recombinant antibody format) on one of its coat proteins (e.g. pIII), with the corresponding variable genes present within its single-stranded circular genome (McCafferty *et al.,* 1990). Filamentous phage are non-lytic, and infect bacteria through the F pilus responsible for conjugal plasmid transfer. Infection is achieved through the binding of the F pilus with the N2 domain of the phage pIII coat protein (Barbas III *et al.,* 2004). The scFv is displayed on this pIII protein. The remaining coat proteins used to make up the tubular phage body are pVI, pVII, pVIII and pIX (see Figure 1.4) (Conroy *et al.,* 2009).

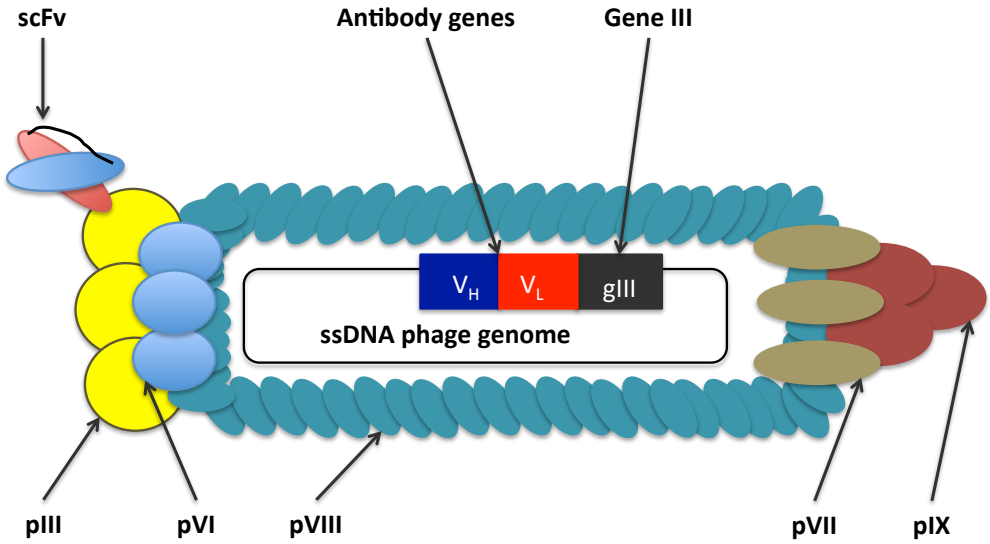

Figure 1.4 – Diagram of an scFv-displaying phage. The scFv is fused to the pIII coat protein; the genes coding the two variable regions of the antibody are located upstream of the gIII gene responsible for coding the pIII protein, providing the phenotype-genotype link.

Display of a scFv on a phage surface is achieved using phage and phagemid vectors. Phage vectors contain the genes required for phage particle assembly and a phage origin of replication. Phagemid vectors contain both a bacterial origin of replication and a phage origin of replication. Antibiotic resistance genes, affinity purification tags and promoter regions for controlled expression are also typical in phagemid vectors. The major advantages offered by phagemid over phage vectors are high transformation efficiencies due to their small plasmid size and soluble antibody expression by the placement of an amber stop codon mutation after the antibody gene and before the geneIII coat gene (Conroy *et al.,* 2009). Phagemid vectors are not, however, capable of producing the packaging proteins required for ssDNA encapsulation by the phage particle during replication. For this, the bacterial cell, already containing the phage ssDNA, must be superinfected by helper phage (for example M13K07 or VCSM13).

Helper phage provides the enzymes and packaging proteins necessary for phage propagation yet have a defective origin of replication, meaning that the phagemid DNA will be packaged over the helper phage DNA (Barbas III *et al.,* 2004). Following transformation of a phagemid vector library into bacterial host cells, the population is superinfected with helper phage and a phage library displaying the antibody repertoire is propagated. To isolate antibodies of interest from the pool, a selective pressure is then applied through panning.

Panning applies a selective pressure to a library-displaying phage population with the aim of isolating high affinity binders against the antigen of interest. Typically, this is achieved through the immobilization of the antigen-of-interest to a solid support (immunotube or 96 well plate) or in solution (conjugated to a magnetic bead). The panning approach utilised in this research was whole cell panning, with PCa cells selected as the antigen. During panning, the antibody-displayed phage library is incubated in the presence of the immobilised antigen to allow binding. The stronger binders are separated from the weaker binders through washing steps and the phage that remain bound are eluted. Elution is usually achieved by altering the pH or with a protease, such as trypsin, which cleaves the antibody from the phage. Eluted phage are then used to reinfect a population of *E. coli* in a rescue step. Following rescue, the infected *E. coli* are cultured and phage propagation is induced. This process can be repeated by precipitating the resultant phage population to start a new round of panning. The stringency of the wash step, applied to remove non-specific binders, is typically increased with each round. This selects for the higher-affinity binders (Barbas III *et al.,* 2004). Panning is repeated until antigen-specific enrichment is obtained.

Following enrichment, the enriched scFv-phage are used to infect a bacterial strain that will recognise the amber stop codon of the phagemid vector (present between the antibody and pIII genes), allowing for soluble expression of the antibody into the periplasmic space of the bacterium without the phage attached. The phagemid vector used in this research is the pComb3X vector (see Figure 2.1).

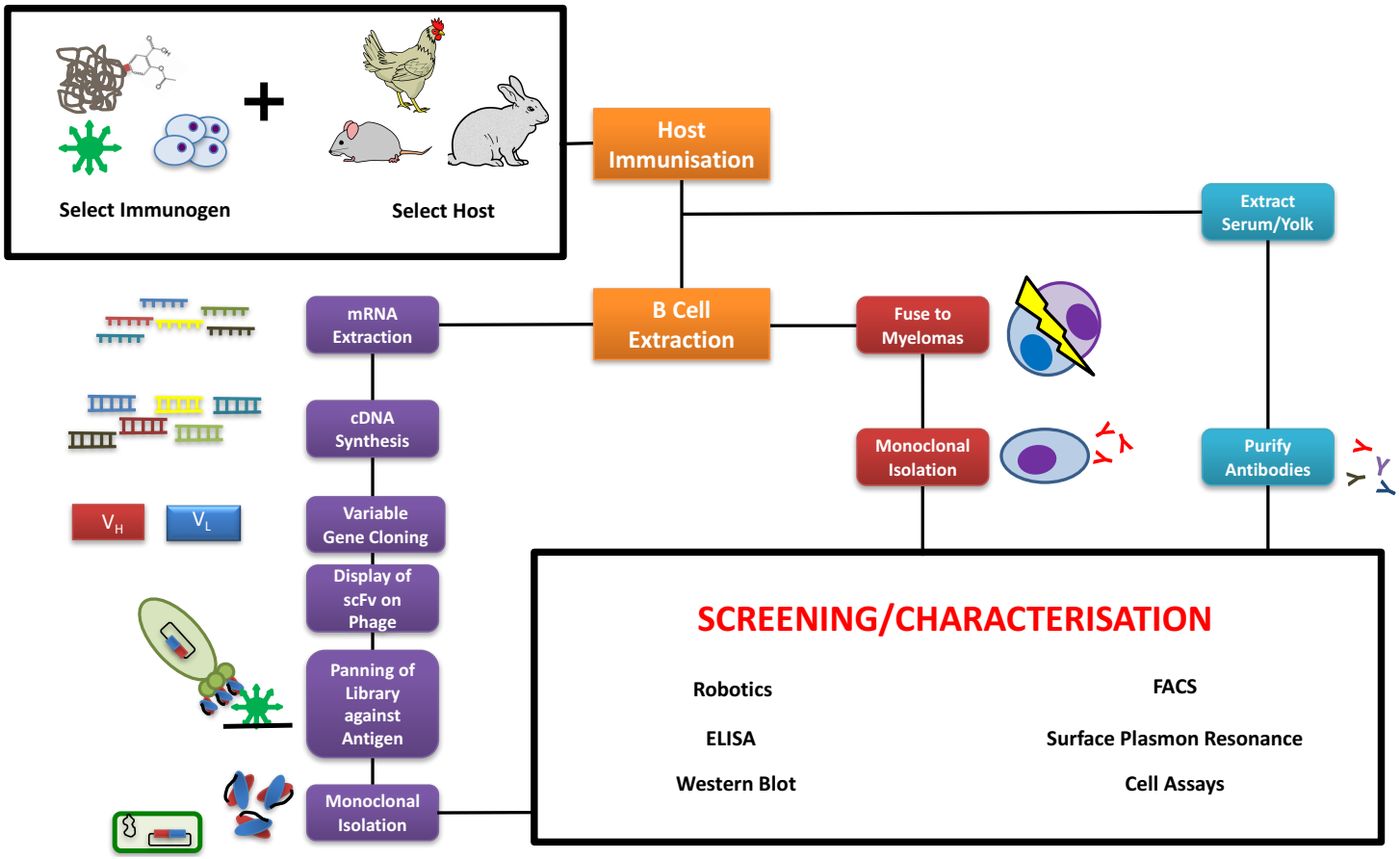

Figure 1.5 - Summary of antibody generation methods following host selection and antigen immunisation. Three routes covered are polyclonal antibody generation (blue flow path), antibody generation using hybridoma technology (red flow path) and antibody generation using recombinant antibody technology (purple flow path).

1.2 The prostate

The prostate is a walnut-sized organ located above the base of the penis and below the bladder, backing onto the rectal wall (see Figure 1.6). It plays a significant role in the male reproductive system, primarily to produce the alkaline seminal fluid responsible for protecting pH-sensitive sperm from the acidic conditions of the vagina. Additionally, it produces Prostate Specific Antigen (PSA), which liquefies the seminal fluid. By contracting during sexual climax, the prostate serves both to expel ejaculate and prevent urine accessing the urethra from the bladder (National Center for Biotechnology Information, 2012).

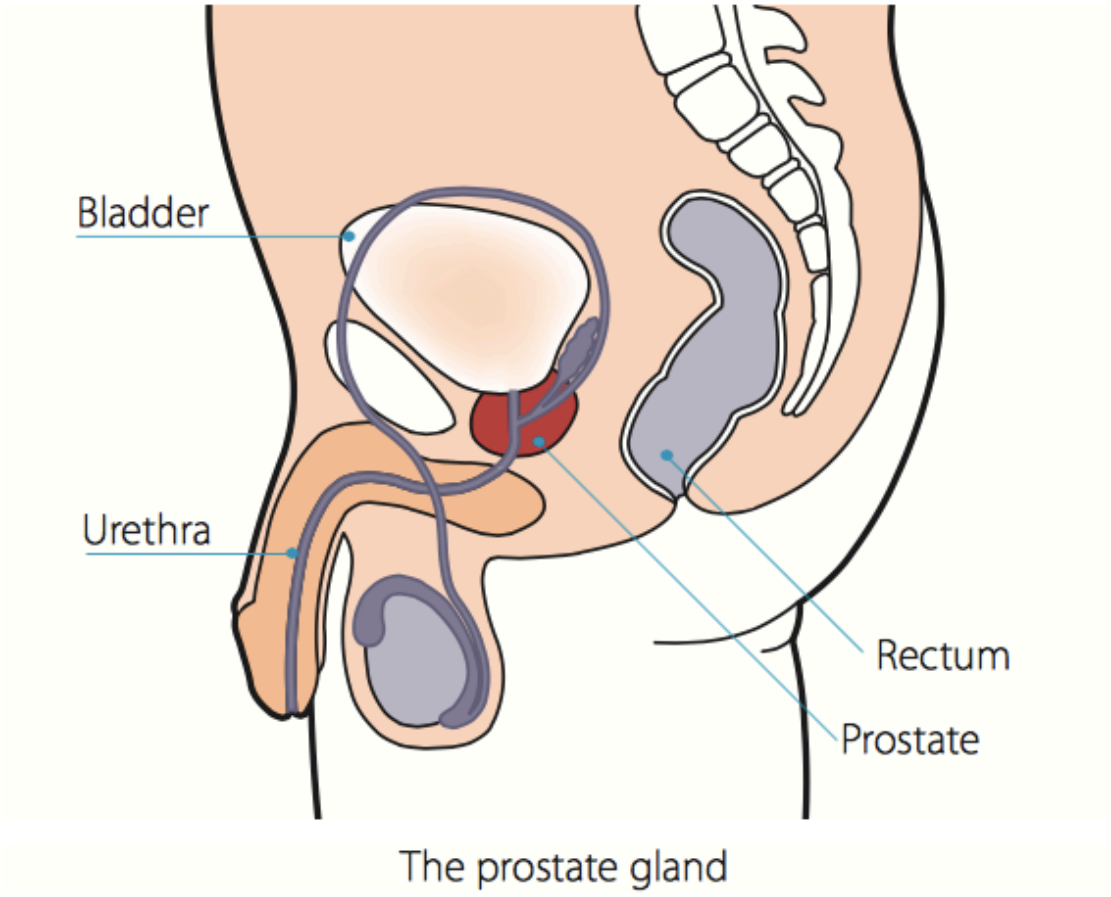

Figure 1.6 – The prostate. The prostate is a walnut-sized organ located beneath the bladder and has a role in the male reproductive system. Diagram obtained from http://www.cancer.ie/sites/default/files/content-attachments/understanding_early_prostate_cancer.pdf [Last accessed 30 April 2017].

1.2.1 Cancer of the prostate

Prostate cancer (PCa) is the most commonly diagnosed form of cancer in men in the Western World, with incidence rates as high as 21% of the male population in European countries (Heidenreich *et al.*, 2014a). In Ireland, 3,000 men are diagnosed with the disease annually (Irish Cancer Society, 2015). Only Lung and Colon cancer have a higher mortality rate. The USA was estimated to have 220,000 new diagnoses in the year 2015 and 27,000 cases of PCa-related death (Siegel *et al.,* 2015).

PCa occurs where cells originating from the prostate proliferate at an abnormal rate to form a tumour. Cancer can remain localized in the prostate as a primary tumour

potentially disrupting the organ's structure or function, or may metastasize to other regions of the body by breaking away from the primary tumour and entering the bloodstream. Common sites of metastasis are the bones and lymph nodes (Karantanos *et al.,* 2015). The majority (~95%) of tumours arising from the prostate are adenocarcinomas, originating from the epithelial layer. The remainder are either sarcomas or neuroendocrine tumours. While the early-stage PCa is often asymptomatic, symptoms associated with the disease include back pain, the presence of blood in urine, erectile dysfunction and difficulty with urination (Weinstein Dunn and Wallace Kazer, 2011).

1.2.2 Risk factors

Although the specific causes of PCa are unknown, there are risk factors that can increase the likelihood of getting the disease. The largest risk factor is age, with a chance of developing the disease increasing from 1 in 10,000 at <40 years of age to 1 in 8 at >60 years of age (Swallow *et al.,* 2012). Race also has an influence on the likelihood of developing the disease. African-Americans are a higher-risk group than other races and incidence is lowest in Chinese men (Haas *et al.,* 2008; Jalloh *et al.*, 2014). Demographic-based diet is considered to be a contributing factor to race-based risk. Countries with a higher intake of dairy-based products and red meat, such as the United States, tend to have a higher incidence rate than countries in East Asia. A study on Asian-Americans showed that their risk of developing the disease roughly doubled based on whether they were born in Asia or the United States. Yet, in either case, incidence was still lower than American Caucasian men, indicating that both race and environment affect risk (Cook *et al.,* 1999). Other dietary factors are also noteworthy, for example, high levels of lycopene in the blood (a result of high tomato intake) causes a reduction in oxidative

genomic damage to prostate cells and an overall reduction in PCa risk (Gann *et al.,* 1999; Chen *et al.*, 2001). High levels of Vitamins E and D, and the mineral selenium, have also been highlighted as potential dietary nutrients that contribute to the prevention of PCa (Schulman and Ekane, 2001).

PCa can be hereditary, with the familial form of the disease likely to develop at an earlier age than the sporadic form. The familial form is characterised by three successive generations developing the disease, three individuals developing the disease in the same nuclear family or two individuals from the same family developing the disease at an age younger than 55 (Swallow *et al.,* 2012). Function-limiting mutations to tumour-suppressor genes, such as the PIK3/Akt survival pathway-regulating phosphatase and tensin homolog gene, have been linked to PCa metastasis and androgen-independence (Phin *et al.,* 2013). Other familial genes attributable to an increased risk of prostate carcinogenesis are the Macrophage-Scavenger Receptor 1 gene and *RNASEL* which is implicated in interferon-mediated antiviral RNA. Additional genes, including the hereditary prostate cancer gene, have been implicated, but not all studies are in agreement on their relevance to the disease (Nelson *et al.*, 2003).

Finally, the general health status and medications being taken by an individual can affect their likelihood of developing PCa. Testosterone, and its metabolite dihydrotestosterone, are important regulators of the size and function of the prostate (Carson and Rittmaster, 2003). PCa is rare in castrated males (Swallow *et al.,* 2012) and studies investigating the effect of 5α-reductase inhibitors, which stop the conversion of testosterone to dihydrotestosterone, show they cause reduction in the risk of developing PCa in addition to a reduction in prostate size (Andriole *et al.,* 2010).

Methods that reduce the amount of free testosterone in the blood, such as physical activity, have been suggested to reduce the risk of PCa additionally. Obesity, on the other hand, has been identified as a significant risk factor for developing PCa (Hsing and Chokkalingam, 2006). Anti-inflammatory drugs such as aspirin and cholesterol-lowering statins have also been shown to reduce the risk of developing the disease (Jacobs *et al.,* 2005; Bansal *et al.,* 2012; Jespersen *et al.,* 2014).

1.3.3 Diagnosis of prostate cancer

PCa is diagnosed with the following tests: Prostate Specific Antigen (PSA) blood test, Digital Rectal Examination (DRE), and Trans Rectal Ultrasound (TRUS) and biopsy of the Prostate (Irish Cancer Society, 2013). Of these methods, the DRE and the PSA blood test are the primary methods of detecting PCa whereas a biopsy is used to confirm the disease state. However, the current PSA-reliant approach for diagnosis of PCa is significantly flawed due to the prevalence of false diagnoses and overtreatment (Gilgunn *et al.*, 2013). In the DRE, a physician feels the prostate through the rectum with a gloved finger. Using this method, the prostate is checked for abnormalities based on its size or structure. The advantages of this examination are that it is inexpensive and is capable of detecting early-stage aggressive cancer that may be missed using PSA testing (Okotie *et al.,* 2007). Yet there are considerable limitations to the DRE test: only the posterior of the gland is checked, detection sensitivity will vary from clinician to clinician and there is a reluctance for patients to undertake the test, due to physical and emotional discomfort. The PSA test quantifies the concentration of PSA in the patient's serum. PSA is produced by the prostate gland to liquefy the semen, granting sperm mobility. It can be found in a man's serum, having diffused into veins from the extracellular fluid. When

the architecture of the prostate is disrupted, higher levels of PSA leak into the blood. Higher levels of PSA are not limited to being a symptom of a prostatic tumour, but are also a symptom of benign prostatic hyperplasia. Hence further, more invasive, tests are required before a patient can be diagnosed with the disease. The range of circulating PSA in the blood is <0.1 ng/mL in a healthy individual to 104 ng/mL in a PCa patient with advanced metastasised disease (Lilja *et al.*, 2008). Currently, PSA serum levels of >4ng/mL warrant further investigation. However 50%-80% of cancer patients identified from this test have an indolent form of the disease, yet may undergo treatments that have strong impacts on their quality of life, raising questions regarding the assay's value (Gilgunn *et al.,* 2013).

If results from the PSA test or DRE are indicative of PCa, the protocol is to biopsy the patient via TRUS for confirmation of the disease (Heidenreich *et al.,* 2014a). TRUS uses ultra sound waves to build a picture of the prostate, to allow for ~12 cores of tissue to be precisely extracted. These samples undergo histological examination and are given a Gleason score. Gleason scoring is used to determine whether a patient has PCa. This system characterizes PCa biopsies based on aggressiveness. Each biopsy is given two scores based on the two most prominent architectural growth patterns in the sample. Values are between 1 and 5, with 1 being closest to resembling healthy prostate tissue. The two scores are added together to give a value between 2-10 per biopsy. The higher the score, the worse the prognosis (Humphrey, 2004). Patients with a score of 6 are considered 'low-risk', those with a score of 7 are considered 'intermediate risk' and those with a score of 8 or higher are considered 'high-risk' (Heidenreich *et al.,* 2014a). There are pitfalls to the Gleason scoring system. Categorising PCa based histologic patterns introduces 'grey-zones' of grading. This negatively affects the reproducibility of

technique (Shah, 2009). Therefore, immunohistochemical (IHC) markers are often used to aid the diagnosis of PCa. There are several markers used and these can be divided into negative markers of malignancy, positive markers of malignancy as well as differential markers. Negative markers of malignancy are typically basal cell markers and include p63, cytokeratin 5/6 and high molecular weight cytokeratin. Positive markers include Alpha-methylacyl-CoA racemase and the Ets-related oncogene product. Differential markers are used to predict the likelihood of secondary tumours as well as the original origin of the tumour – for example prostein, PSA and Prostate Specific Membrane Antigen are indicators of prostatic tissue. However, considerable pitfalls remain for IHC-based diagnosis of cancer. For example, the commonly-used malignancy marker Alpha-methylacyl-CoA racemace is absent in up to 25% PCa cases, yet may be found in non-malignant prostatic intraepithelial neoplasia (Paner *et al.*, 2008; Brimo and Epstein, 2012). Hence, reliable diagnoses rely on panels of biomarkers, in combination with the inputs of an experienced physician. If Gleason scores, PSA values and DRE evaluation are indicative of PCa, staging tests can be carried out to assess the likelihood that the cancer has spread to other sites in the body. Staging tests include biopsies of the seminal vesicle and nearby lymph nodes, to check for cancer cells. Imaging (Magnetic Resonance Imaging or CAT scan) of the patients tissue can be carried out to search for tumours that have metastasized (National Cancer Institute, 2014). Bone scans are also carried out, as bone metastases are found in 90% of patients that have died from PCa (Keller, 2007).

Staging can be either clinical (based on a PSA, DRE and imaging) or pathological (based on tissue analysis and the TNM system) (Weinstein Dunn and Wallace Kazer, 2011). The TNM system characterizes cancer by giving it a score based on the size and extent of the

primary tumour (T1-4), whether or not cancer cells have been found in the lymph nodes (N0-1) or whether it has metastasized to other regions of the body (M0-1). Staging can range from I to IV, with an increase in stage representing a more advanced form of the disease (Heidenreich *et al.,* 2014a). The stage of the cancer determines the optimum treatment to undertake. The TNM system is outlined in Table 1.1.

In summary, current methods to detect PCa rely on the limited PSA assay and DRE. For a more definitive diagnosis, patients must undergo a stressful and painful biopsy of the prostate. Diagnosis of a tumour through biopsy requires an experienced clinician and a panel of imperfect biomarkers. If PCa is confirmed, all of the possible investigative steps undertaken are then considered before determining the preferred line of treatment. The discovery of novel biomarkers to PCa, that could be incorporated into this diagnosis process, may improve diagnosis of the disease and have a positive impact on the determination of the best treatment options.

Table 1.1 - PCa staging guide (National Cancer Institute, 2015).

Group	T	N	M	PSA (ng/mL)	Gleason score
I	T1a-c	0	0	<10	≤6
	T2a	0	0	<10	≤6
	T1-2a	0	0	X	X
IIA	T1a-c	0	0	<20	7
	T1a-c	0	0	≥10<20	≤6
	T2a	0	0	≥10<20	≤6
	T2a	0	0	<20	7
	T2b	0	0	<20	7
	T2b	0	0	X	X
IIB	T2c	0	0	Any	Any
	T1-2	0	0	≥20	Any
	T1-2	0	0	Any	≥8
III	T3a-b	0	0	Any	Any
IV	T4	N0	Any	Any	Any
	Any	N1	Any	Any	Any
	Any	Any	1	Any	Any

The 'T' column describes the primary tumour. T1 – tumour was not visible by imaging. T1a – tumour was found by accident. T1b – tumour was found by accident but is present in >5% of tissue removed. T1c – tumour was found by biopsy following raised PSA levels. T2 – tumour is confined to the prostate. T2a – tumour is confined to one side of the prostate. T2b – tumour confined to one side of the prostate but exceeds the midsection of prostate. T2c – tumour is in both sides of the prostate. T3 – tumour extends through the prostate. T3a – tumour has not spread to the seminal vesicles. T3b –tumour has spread to the seminal vesicles. T4 – tumour has invaded adjacent tissues. The 'N' column gives the status of the regional lymph node, with 0 considered negative and 1 considered positive for metastasis. The 'M' column refers to metastasis to tissues outside the

prostate with 0 considered negative and 1 considered positive. The Gleason score is given as a value between 2-10. 'X' in any column is considered as 'data not accessible'. 'Any' in any column is considered as any value possible.

1.2.4 Current treatment options for prostate cancer

For treating PCa, the European Association of Urology has established guidelines. These vary based on the stage of the disease, ranging from a primary low-risk tumour (Heidenreich *et al.,* 2014a) to a more advanced hormone-independent stage (Heidenreich *et al.,* 2014b). The following treatment options are based on these recommendations.

For patients of >70 years of age with a low-risk PCa, active surveillance is considered the best treatment. Stage I/II PCa is considered very low-risk for these patients, as their life expectancy is shorter than the time it would take for the PCa to progress to a stage that poses a significant health risk. With active surveillance the cancer is continually monitored through DRE and PSA tests in case of progression, after which more invasive treatments are considered (Weinstein Dunn and Wallace Kazer, 2011). Radical prostatectomy is the only surgical option available for patients with localized PCa. Due to its invasiveness, it is only considered beneficial to patients with low or intermediate risk PCa (typically Stage I or IIA) and a life expectancy of >15 years. The surgery removes the prostate gland and hence any primary tumours present. Sexual and urinary tract complications are common with this procedure, particularly erectile dysfunction and urinary incontinence although these can be post-operatively treated. The negative side effects associated with this procedure can be minimized, provided the nearby nerve endings and tissues are undamaged during surgery.

An alternative option for patients who do not wish to undergo surgical intervention is radiation therapy. External beam radiation therapy uses a directed external source of radiation to focus on the location of the tumour and reduce collateral damage to adjacent tissues. Brachytherapy places radiation-emitting seeds in or around the tumour and can be temporary or permanent depending on the level of radiation they emit. Brachytherapy can be used on its own or in combination with External beam radiation therapy. These treatments offer similar success rates to radical prostatectomy with recurrence-free survival as high as 83% of patients after ten years. However, they also have the drawbacks of sexual dysfunction and urinary incontinence (Weinstein Dunn & Wallace Kazer 2011; Swallow *et al.,* 2012; Heidenreich *et al.,* 2014b).

For treating high-risk, localized PCa there is no agreed best treatment approach. Therapy is carried out on a case-by-case manner. If the tumour is of a small volume, radical prostatectomy is a viable first measure; however, multiple treatments are usually required as there is a significant likelihood of nodal invasion. Even without nodal invasion a combination of radiotherapy and hormone therapy has been recommended (Mottet *et al.,* 2015). Over half of patients diagnosed with advanced PCa will require salvage radiotherapy (following an ineffective primary therapy), adjuvant therapy or hormonal therapy (Heidenreich *et al.,* 2014b). Hormone therapy targets androgens, the hormones responsible for prostate cell growth and regulation. The main androgens are testosterone and dihydrotestosterone (DHT). Testosterone is produced primarily by the testes but also in the adrenal glands whereas DHT is produced primarily in the stroma of the prostate via enzymatic conversion of free testosterone using 5α-reductase. Androgens bind the androgen receptor, located in the cytoplasm of prostate epithelial

and PCa cells, inducing changes leading to cell proliferation (Harris *et al.,* 2009). Reducing the level of circulating androgens in the blood can halt and shrink the growth of PCa, at least temporarily. Castration is achieved either chemically or surgically. Chemically-induced castration is carried out using agonists of the luteinizing hormone-releasing hormone. This hormone causes the release of luteinizing hormone from the pituitary gland, inducing testosterone production in the testes. Androgen production should resume following cessation of the treatment. Surgical castration is achieved with an orchidectomy and cannot be reversed. The major side-effect of hormone deprivation is sexual dysfunction and the efficacy of androgen deprivation therapy varies based on the stage of the PCa. In patients with metastatic cancer, the disease can be halted for 2-3 years before castration resistance develops – after which the survival time is around 18 months (Harris *et al.,* 2009).

Castration-resistant PCa occurs when the PCa cells proliferate despite androgen deprivation therapy. How this is achieved varies with the cancer, however, it will typically involve mutations to the androgen receptor responsible for proliferation or an increase in androgens in the tumour microenvironment (Karantanos *et al.,* 2015). The approved treatments for castration-resistant PCa are docetaxel, cabazitaxel, sipuleucel-T, abiraterone acetate plus prednisone and enzalutamide (Heidenreich *et al.,* 2014b). Docetaxel administered with the immunosuppressant prednisone is currently the treatment of choice and can extend a patient's life an average of 19 months (Berthold *et al.,* 2008). Docetaxel prevents microtubule disassembly, which is a requirement when a cell is undergoing mitosis. Second-line treatments are available in case of the development of docetaxel-resistance in the disease. These include: Cabazitaxel, a tubulin-binding taxane, with prednisone; Abiraterone acetate, an aromatase inhibitor

that targets androgen biosysnthesis, plus prednisone; Enzalutamide, an androgen receptor antagonist and Sipuleucel-T, an autologous vaccine in which a patient's white blood cells are removed from the blood and fused with prostatic acid phosphatase (a common PCa antigen) before re-entry into the blood as antigen-presenting cells. These treatments, however, will only prolong a patient's life by a matter of months and have significant toxicity-based side effects (Kantoff *et al.*, 2010; Scher *et al.*, 2012; Heidenreich *et al.*, 2014a, 2014b). A common occurrence (>90%) with castration-resistant PCa is bone metastases. Tumour formation on the bone leads to a rapid decrease in quality of life, as well as morbidity. Treatments shown to increase survival rates and improve the quality of life of patients with bone metastases are the osteoclast-limiting zoledronic acid (osteoclast or bone break-down activity is elevated in PCa bone lesions), the calcium mimicking Radium 223 (which delivers a contained radioactive payload to areas of bone growth – typical with bone metastases) and the monoclonal antibody, Denosumab (which inhibits osteoclast formation by targeting RANKL on their surface) (Saad et al., 2004; Fizazi et al., 2011; Parker et al., 2013).

1.3. Antibodies as tools against cancer

1.3.1 Current applications of antibodies against cancer

With the development of hybridoma technology, the use of antibodies in the medical field has become routine procedure. Antibodies have been engineered to suit a range of diagnostic purposes from at-home pregnancy tests, to rapid detection of a disease state in the clinic, to detailed disease prognosis in a lab (Kijanka *et al.*, 2010; Lemass *et al.*, 2014; Sharma *et al.*, 2015). In addition to their use as diagnostic and prognostic tools, antibodies have become established as one of the most successful therapeutic strategies in both oncological and non-cancer treatments. Due to their high affinity for a target antigen, they offer the advantage of localizing to the site-of-interest in comparison to standard chemotherapeutics, thereby reducing the side effects associated with cancer treatments.

Therapeutic antibodies typically target an antigen that is upregulated in a cancerous disease state that plays a role in cell survival and proliferation. When a cell becomes carcinogenic, its antigen marker profile is altered. Following identification of a target antigen, a mAb to that antigen may deliver its therapeutic effect via a number of mechanisms: inhibition of cell signalling, induction of apoptosis, antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity and delivery of a toxic payload to tumour cells (these approaches are illustrated in Figure 1.7) (Glassman and Balthasar, 2014). A limitation of such a targeted approach is the reliance on the disease state retaining the antigen-of-interest, with mutations to the antigen potentially resulting in acquired resistance to monotherapy. Consequently, combination therapies are also being explored for improved clinical outcomes (Sharma *et al.,* 2017). A list of the

therapeutic antibodies and their associated disease, currently approved for use is available in Table 1.2.

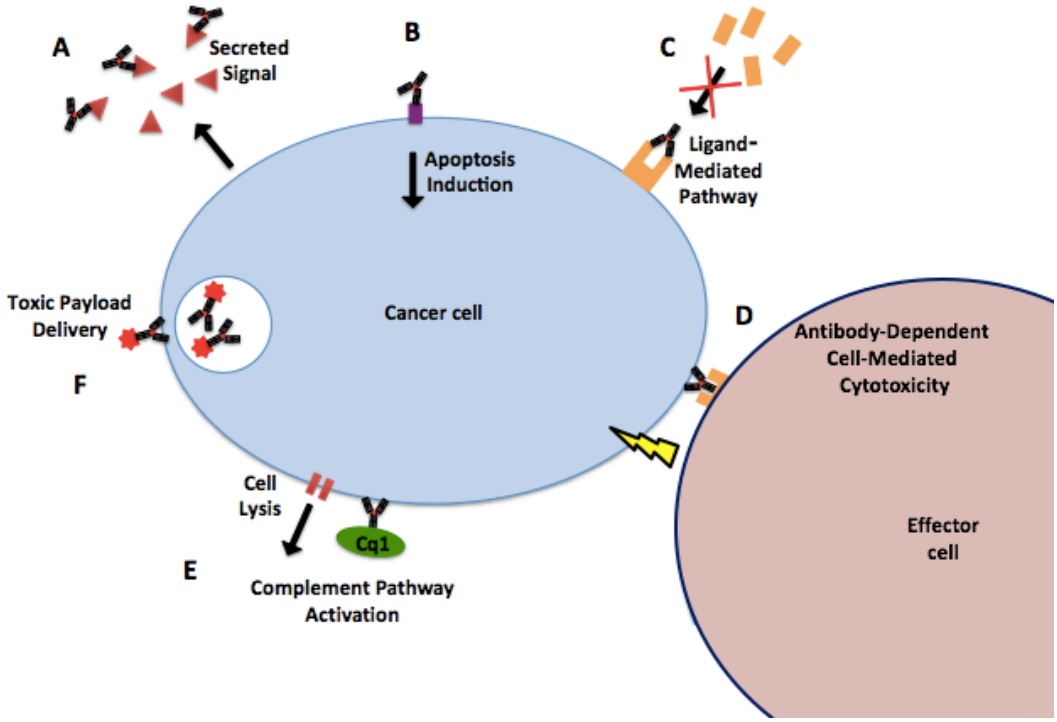

Figure 1.7 - Antibody-mediated mechanisms of treating cancer. A – Antagonist of secreted proliferation-inducing ligand. B – Induction of apoptosis through surface receptor. C – Antagonist of ligand-activated proliferation receptor. D – Antibody-dependent cell-mediated cytotoxicity. E - Complement pathway activation. F – Antibody-directed payload delivery.

Table 1.2 - Antibody treatments approved or in review for use in both the European Union and United States.

Name	Target	Format	Disease	EU Approval	US Approval
Alirocumab	PCSK9	Human IgG1	High cholesterol	In review	In review
Mepolizumab	IL-5	Humanized IgG1	Severe eosinophilic asthma	In review	In review
Necitumumab	EGFR	Human IgG1	Non-small cell lung cancer	In review	In review
Nivolumab	PD1	Human IgG4	Melanoma	In review	2014
Dinutuximab	GD2	Chimeric IgG1	Neuroblastoma	In review	NA
Secukinumab	IL-17a	Human IgG1	Psoriasis	Decision pending	In review
Evolocumab	PCSK9	Human IgG2	High cholesterol	In review	In review
Blinatumomab	CD19, CD3	Murine bispecific tandem scFv	Acute lymphoblastic leukemia	In review	2014
Pembrolizumab	PD1	Humanized IgG4	Melanoma	In review	2014
Ramucirumab	VEGFR2	Human IgG1	Gastric cancer	Decision pending	2014
Vedolizumab	α4β7 integrin	humanized IgG1	Ulcerative colitis, Crohn disease	2014	2014
Siltuximab	IL-6	Chimeric IgG1	Castleman disease	2014	2014

Obinutuzumab	CD20	Humanized IgG1	Chronic lymphocytic leukemia	2014	2013
Ado-trastuzumab emtansine	HER2	humanized IgG1 (immuno-conjugate)	Breast cancer	2013	2013
Raxibacumab	*Bacillus anthracis* PA	Human IgG1	Anthrax infection	NA	2012
Pertuzumab	HER2	humanized IgG1	Breast Cancer	2013	2012
Brentuximab vedotin	CD30	Chimeric IgG1 (immunoconjugate)	Hodgkin lymphoma, systemic anaplastic large cell lymphoma	2012	2011
Belimumab	BLyS	Human IgG1	Systemic lupus erythematosus	2011	2011
Ipilimumab	CTLA-4	Human IgG1	Metastatic melanoma	2011	2011
Denosumab	RANK-L	Human IgG2	Bone Loss	2010	2010
Tocilizumab	IL6R	Humanized IgG1	Rheumatoid arthritis	2009	2010
Ofatumumab	CD20	Human IgG1	Chronic lymphocytic leukemia	2010	2009
Canakinumab	IL1b	Human IgG1	Muckle-Wells syndrome	2009	2009
Golimumab	TNF	Human IgG1	Rheumatoid and psoriatic arthritis, ankylosing spondylitis	2009	2009
Ustekinumab	IL12/23	Human IgG1	Psoriasis	2009	2009

Certolizumab pegol	TNF	Humanized Fab (pegylated)	Crohn disease	2009	2008
Catumaxomab	EPCAM/ CD3	Rat/mouse bispecific mAb	Malignant ascites	2009	NA
Eculizumab	C5	Humanized IgG2/4	Paroxysmal nocturnal hemoglobinuria	2007	2007
Ranibizumab	VEGF	Humanized IgG1 Fab	Macular degeneration	2007	2006
Panitumumab	EGFR	Human IgG2	Colorectal cancer	2007	2006
Natalizumab	A4 integrin	Humanized IgG4	Multiple sclerosis	2006	2004
Bevacizumab	VEGF	Humanized IgG1	Colorectal cancer	2005	2004
Cetuximab	EGFR	Chimeric IgG1	Colorectal cancer	2004	2004
Efalizumab	CD11a	Humanized IgG1	Psoriasis	2004#	2003#
Omalizumab	IgE	Humanized IgG1	Asthma	2005	2003
TositumomabI1-1	CD20	Murine IgG2a	Non-Hodgkin lymphoma	NA	2003#
Ibritumomab tiuxetan	CD20	Murine IgG1	Non-Hodgkin lymphoma	2004	2002
Adalimumab	TNF	Human IgG1	Rheumatoid arthritis	2003	2002
Alemtuzumab	CD52	Humanized IgG1	Chronic myeloid leukemia#, multiple sclerosis	2001#, 2013	2001#, 2014
Gemtuzumab ozogamicin	CD33	Humanized IgG4	Acute myeloid leukemia	NA	2000#
Trastuzumab	HER2	Humanized IgG1	Breast cancer	2000	1998
Infliximab	TNF	Chimeric IgG1	Crohn disease	1999	1998

Palivizumab	RSV	Humanized IgG1	Prevention of respiratory syncytial virus infection	1999	1998
Basiliximab	IL2R	Chimeric IgG1	Prevention of kidney transplant rejection	1998	1998
Daclizumab	IL2R	Humanized IgG1	Prevention of kidney transplant rejection	1999#	1997#
Rituximab	CD20	Chimeric IgG1	Non-Hodgkin lymphoma	1998	1997
Abciximab	GPIIb/IIIa	Chimeric IgG1 Fab	Prevention of blood clots in angioplasty	1995*	1994
Muromonab-CD3	CD3	Murine IgG2a	Reversal of kidney transplant rejection	1986*	1986#

Legend: * - Country specific approval. # - Withdrawn since approval. NA – Not approved.

Table taken from (Reichert, 2015).

1.3.2 Discovery of novel cancer targets

A major obstacle in the development of diagnostic and therapeutic antibodies is the identification of novel biomarkers and therapeutic targets. Discovery is a slow and laborious process. The majority of antibodies, particularly those with a therapeutic function, will not make it to market. The cost associated with discovery and development of antibodies is a significant limiting factor in the field (Liu, 2014). A review by Drucker and Krapfenbauer (2013) highlighted that improvements to the development process would be greatly beneficial, and can be achieved as early as the discovery phase. This section describes the current methods used to discover novel antigens.

1.3.2.1 Classical methods

The earliest methods of antibody development rely on identification of a target, followed by generation of an antibody to that target using hybridoma technology. An excellent example of this was the development of transtuzumab, or its commercial name, Herceptin. Following the discovery of the Her2/neu receptor in 1976 and confirmation of overexpression in breast cancer in 1985, hybridoma technology was used to produce an antibody to it in 1987. After demonstration of its therapeutic potential, the antibody was humanized and started clinical trials in 1992. It received FDA approval for the treatment of HER2-positive metastatic breast cancer in 1998 (Genentech, 2016). The development of Herceptin relied on DNA cloning and immunoassays such as Western blotting and Enzyme-Linked Immunosorbent Assay (ELISA). While this approach is still used today, an associated bottleneck is the required predetermination of the antigen target, before antibody development. Research is

ongoing into adapting the fundamentals of this process to the discovery of unknown antigens. A newer approach is to screen hybridomas for antibodies that induce a therapeutic effect upon binding, such as induction of apoptosis, rather than binders to a predetermined antigen. The bottleneck of this approach is the feasibility of screening vast numbers of clones manually. Both robotics and microfluidics have been employed to enhance this screening process (Abdiche *et al.*, 2014; Fitzgerald *et al.*, 2015b).

1.3.2.2 Display technologies

While the majority of antibodies approved for clinical use have been generated using hybridoma technology, phage display has been applied to successfully generate FDA-approved therapeutic antibodies such as Adalimumab (sold under the brand name HUMIRA) (Kim *et al.*, 2011). The traditional disadvantages of phage display technology, such as therapeutic immunogenicity, have been circumvented with optimised humanization protocols (Townsend *et al.*, 2015). It is estimated that up to 35% of mAbs in clinical trials have been generated through phage display technology (Omidfar and Daneshpour, 2015).

In addition to generating antibodies to preselected antigens, phage display can also be used to select for antibodies against unknown antigens. This is achieved through cell-panning (see Figure 1.8). Cell panning uses a cell line as the antigen-of-interest during the panning process. This allows for library enrichment against multiple antigens present in the cell line. The process can select for antibodies to internal antigens, in addition to those on the cell surface (Zhou *et al.*, 2012; Keller *et al.*, 2015). Typically libraries that undergo panning have been generated in a host immunized with a cell line (Hoogenboom *et al.*, 1999), yet the process is powerful enough to successfully pan

antibodies-of-values from a naïve library (Keller *et al.*, 2015). This technology has been acknowledged to be useful in the generation of antibodies to non-predetermined antigens for drug discovery purposes (Hoogenboom, 2005; Even-Desrumeaux *et al.*, 2013). There is, however, a major drawback to the technology. That is, the inability to distinguish antigens-of-interest from ubiquitous antigens. Consequently, antibodies selected using cell panning require extensive characterisation to determine the identity of their corresponding antigens. The standard approach to achieve this utilises immunoassays and bioassays as a primary screening tool for antibodies of interest in a 'function-before-identity' strategy. The antigens corresponding to the antibodies-of-interest are typically identified through co-immunoprecipitation with the antibody from cell lysate followed by mass spectrometry to precisely confirm the identity of the target antigen. The incorporation of a depletive panning step, against a phenotypically distant cell line, is an often-used addition to the process that aims to reduce the likelihood of isolating antibodies to ubiquitous antigens. However, employing a negative selection step risks the loss of antibodies to antigens that have a modified expression in the disease state and is better-suited for identifying antigens that are simply present or absent in the disease state. The drawbacks of negative-selection mean it has not been universally adopted in the field (Hoogenboom *et al.*, 1999; Siva *et al.*, 2008; Keller *et al.*, 2015). Whether depletive panning is incorporated or not, the downstream steps involved in antigen identification remain rigorous, and not all publications include them (Popkov *et al.*, 2004). Until improvements to the efficiency of antigen discovery via cell panning are made, the technology is likely to remain less favourable than more targeted methods.

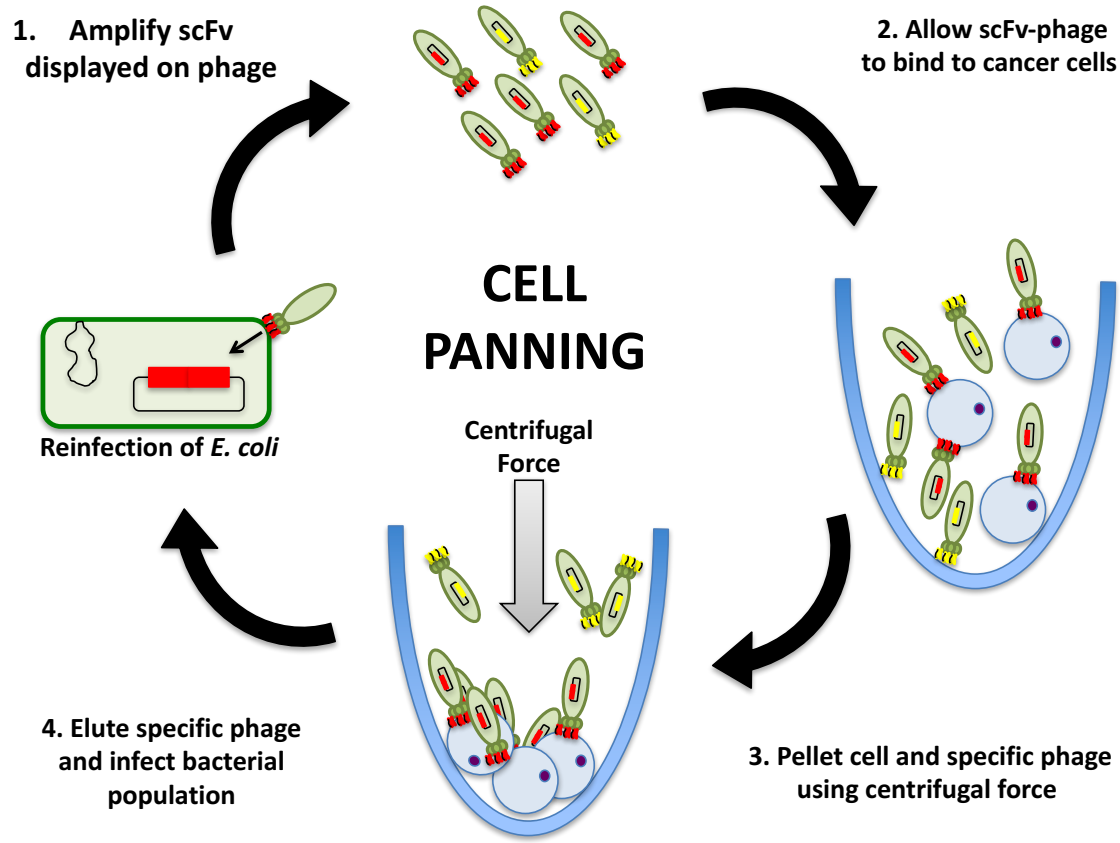

Figure 1.8 – Diagram of cell-based panning, where PCa cells are the target antigen. 1 – an antibody-displaying phage library is generated using helper phage super infection of a phagemid library. 2 – the antibody-displaying phage are incubated with live PCa cells to allow specific binding of phage. Note: cell-specific antibodies are represented in red, non-specific are represented in yellow. 3 – the PCa cells are pelleted through centrifugation, with phage bound to the cells also pelleted. Non-specific phage remain in the supernatant, as they fail to bind and are then removed. 4 – The specific phage are eluted from the pelleted cells by proteinase activity or pH alteration. The specific phage are then removed, used to infect a bacterial population and propagated *via* helper phage. The propagated phage undergo the additional rounds of the same panning procedure. Multiple repeats of this procedure will enrich for specific phage until monoclonal scFv-expressing cells are isolated and characterised.

1.4.2.3 Mass spectrometry-based proteomics

Comparing the protein profiles of a cancerous disease state and a healthy disease state is an established approach for the identification of novel cancer antigens. Protein profiles may be acquired from several sources, including sera and tissue from patients,

or cell lines. Proteins that are differentially expressed between disease and heathy states are considered as potential biomarkers and therapeutic targets. Another approach to biomarker discovery investigates differences in the autoantibody profiles between diseased and healthy individuals. Autoantibodies are antibodies generated by an individual against self-proteins.

Two-dimensional gel electrophoresis sequentially applies two perpendicular fields of charge to a complex protein mixture, separating them based on their isoelectric point and mass. Lysis, separation and comparison of healthy and diseased organ tissue using two-dimensional gel electrophoresis is applied to highlight differentially expressed proteins. Following identification of the potential biomarkers, the protein spots are excised from the gel and identified based in part on their peptide mass using Mass Spectrometry (MS). For autoantibody characterisation, tissue or cell-derived protein profiles, separated by two-dimensional gel electrophoresis, are transferred onto membranes for Western blotting. Investigating for differences in proteins bound by antibodies isolated from the sera of healthy and cancer patients can also highlight biomarkers. A significant obstacle to these approaches is the difficulty to identify minute amounts of differentially-expressed antigen. Antigens may be masked by more abundant proteins, or simply not visualized and, hence, not detected. This is particularly common for antigens of a low molecular weight or isoelectric point (Lu *et al.*, 2007). Improvements to biomarker-identification have been made by comparing profiles using imaging software as well as by replacing manual protein-spot picking with robotics. Proteins that overlap on a gel can also be separated accurately using liquid chromatography before MS. Ultimately, the proteins-of-interest still must be isolated

from the gel, fragmented and identified based on their peptide mass fingerprint using a mass spectrometer.

The introduction of 'isobaric tags, for relative and absolute quantification' technology, has addressed the shortcomings in quantitation of two-dimensional gel electrophoresis and MS (Lu *et al.*, 2007). Proteins from different sources can be labelled with mass-based reporter tags and analysed simultaneously in a mass spectrometer. Identical proteins from separate samples can then be quantified and their origin determined based on their isobaric tags. The power of isobaric tag technology for biomarker discovery was demonstrated when one group compared healthy and cancerous ovarian epithelial tissue, identifying 1,259 unique proteins, of which 205 were differentially expressed (Wang *et al.*, 2012). Overall, MS-based proteomics provides an excellent approach to identifying cancer-associated antigens, yet includes drawbacks such as the high cost and accessibility of the technology, as well the skill required for interpretation of results.

1.3.2.4 Protein array technology

Protein arrays are surfaces hosting a multitude of immobilised proteins, on which immunoassays or other analyses are performed. A primary application is in investigating protein-protein interactions in a potentially high-throughput format. Arrays are constructed by spotting protein libraries onto glass or filter membranes using robotic – based technologies to ensure uniform production. The proteome immobilised onto the array can vary and the formats of immunoassays carried out are quite flexible. Some arrays consist of a characterised protein library immobilized on a membrane, allowing for the identification of the cognate antigen to an uncharacterised antibody (Bussow *et*

al., 1998). There are also arrays where uncharacterised antibodies are spotted onto the surface for screening against an introduced protein-of-interest. A commonly used application is the immobilisation of characterised antibodies onto the array, which is then probed with a complex solution, such as a cell lysate, to determine whether antigens-of-interest are present (El-Haibi *et al.*, 2012). These three approaches are illustrated in Figure 1.9

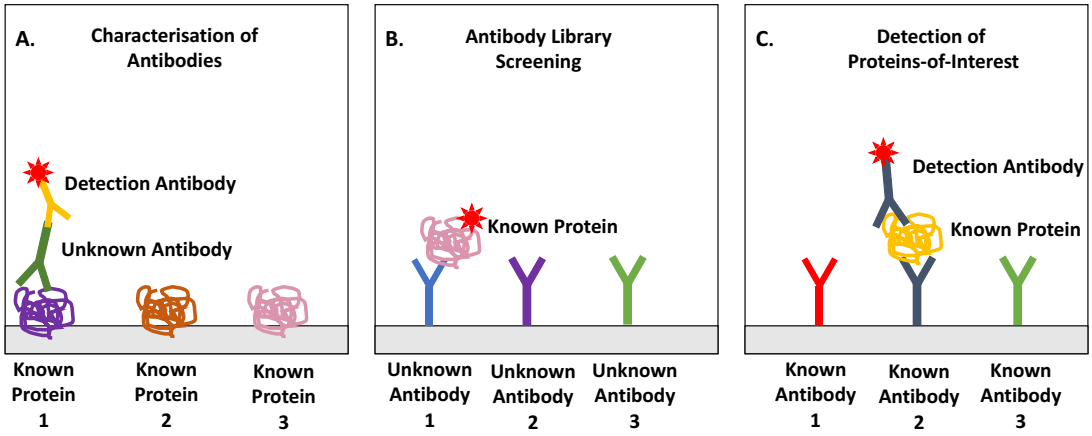

Figure 1.9 - Applications of protein arrays. A – Array consists of characterised proteins. Probing of the array with an uncharacterised antibody reveals its cognate antigen. B – Array consists of uncharacterised antibodies. Probing of the array with a fluorescently tagged protein-of-interest identifies an antibody specific to it. C – Array consists of antibodies specific to proteins of interest. Probing of the array with complex solutions such as cell lysate reveals whether proteins-of-interest are present.

The arrays applied in this research were produced by Bussow *et al.* (1998) and were made up of a characterised protein library. One of the earliest applications of the technology demonstrated the simultaneous identification of antigens to low numbers of scFv, isolated from a naïve human library (Holt *et al.*, 2000). Subsequent proof-of-concept studies demonstrated the technology's ability to identify the antigens of antibodies generated through polyclonal and hybridoma technology (Kijanka *et al.*, 2009a; Kijanka *et al.*, 2009b). Since then, the arrays have been used to identify the specific antigens of autoantibodies present in the sera of cancer patients. The purpose of such investigations was biomarker discovery. Multiple studies have since been carried out, comparing sera from disease positive and negative patients from a range of cancers, including prostate, colorectal and lung, identifying a plethora of biomarkers in the process (Wang *et al.*, 2005; Kijanka *et al.*, 2010; Brezina *et al.*, 2015). Antigens identified through protein array experiments have been clinically validated through follow-up

studies, further demonstrating the potential of the technology (Fitzgerald *et al.*, 2013, 2015a; O'Reilly *et al.*, 2015).

While protein array technology has made strides in the clinical biomarker discovery field, it has shown relatively few breakthroughs in improving antibody generation technology. A notable exception was a recent study carried out by Kibat *et al.,* (2015). Here, antibody array technology was utilised to augment the antibody generation process. Almost 500 scFv were isolated after panning against 134 preselected cancer-associated antigens. These antibodies were then displayed as antibody microarrays. Protein-containing lysates, from pancreatic cancer and healthy tissue, were incubated with the arrays. The binding profiles of each array was analysed for antibodies that specifically bound proteins from cancerous tissue lysate, highlighting the antigens as potential pancreatic cancer biomarkers. The antibodies of interest were then isolated and characterised individually on a protein microarray, made up of 54 oncoproteins, to reveal their cognate antigens. This experimental approach successfully identified potential biomarkers in pancreatic cancer, and facilitated the identification of recombinant antibodies to the identified biomarkers. A slight drawback of the study was that the antigens were preselected – limiting the scope of discovery to already characterised onco-markers.

1.4 Thesis aims

PCa is the most prevalent form of cancer in men in the western world. However, the current methods of detecting and treating the disease are severely limited. Improvements will come following the discovery of superior PCa biomarkers that can be used to detect and discriminate between indolent and aggressive forms of the disease. In addition, the identification of novel therapeutic targets may reduce PCa-related mortalities. Antibody generation technology can provide the diagnostic and therapeutic tools needed for these advances; however, its capacity for biomarker discovery is currently inferior to other established proteomic technologies. Improvements to the efficiency of antibody generation, against antigens that have not been predetermined, may increase the output of antibodies to uncharacterised antigens in the field and improve our current understanding of PCa. This research aimed to address this output limitation by using protein array technology to augment the standard approach to phage display-based biomarker discovery. This was to be achieved by:

- The generation of a recombinant scFv library displayed on phage from chickens, a commonly used host, immunised with PCa cells lines.
- The application of protein array technology to characterise the avian immune response to PCa cells as well as the scFv library generated. The antigens identified were to be examined using bioinformatics to reveal trends in enrichment.
- Enriching the scFv-phage library against unknown antigens present in PCa cells using cell panning and isolating individual antibody clones for characterisation.
- Combining the panning-isolated antibodies and introducing them to a protein array for single-step identification of their cognate antigens. This antigen list was

also to be analysed using bioinformatics, with a view to revealing trends in antigen enrichment.

- The specificity of the secondary antibodies to be used in each profiling experiment will be characterised by the technology prior to use, providing information on their suitability for future immunoassays and in an effort to avoid incorrect identification of potential antigens or biomarkers
- Validate the strategy by expressing a subset of the array-identified scFv antigens and screening the scFv pool to identify and characterise their corresponding binders using ELISA, Western blotting and surface plasmon resonance based technologies.

The successful application of the proposed workflow would reverse the 'function-before-identity' strategy currently utilised in the field of phage display-based biomarker discovery, as well as removing the current requirement for antigen identification via MS (see Figure 1.10).

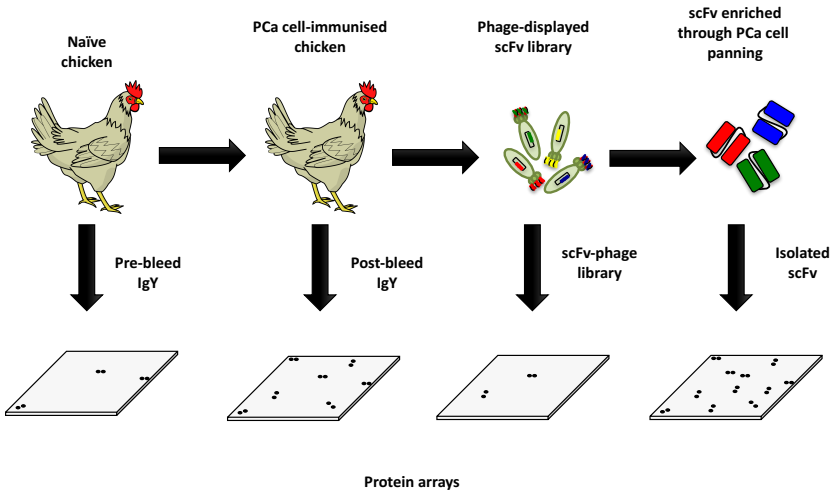

Figure 1.10 - Schematic of the workflow utilised in this research. A chicken will be immunised with a combination of PCa cell lines to induce an immune response. The immune response will be converted to scFv-phage to generate a recombinant scFv-phage library. The library will be panned against PCa cells until enrichment is observed. A panel of scFv-encoding *E. coli* will be isolated for characterisation. Protein array technology will be applied to reveal the antigen profile at each of these steps to further our understanding of the recombinant antibody generation process.

Chapter 2 - Materials and methods

2.1 Materials

2.1.1 Equipment list

Apparatus	Supplier
Nikon Diaphot Inverted Tissue Culture Microscope	Nikon Instruments Inc., 1300 Walt Whitman Road, Melville, New York 11747-3064, United States.
Fuji FLA5100 High Resolution Phosphorimager and Fluorescence Imaging System	Brennan and Company, 61 Birch Avenue, Stillorgan Industrial Park, Stillorgan, Ireland.
Clare Chemical Dark Reader DR195 transilluminator	Clare Chemical Research Inc., 995 Railroad Ave, Unit E Dolores, CO 81323, United States.
Balances (Chyo-Model JK-180 and Mettler Model-PJ300) **Orion 3 star-pH meter** **Thermo Scientific™ Pierce™ G2 Fast Blotter**	Medical Supply Company Ltd., Damastown, Mulhuddart, Dublin 15, Ireland.
Bio-Rad Gene Pulser Xcell	Alpha Technologies,

Geldoc EZ System Image Lab ™ **Bio-Rad PowerPac ™ Basic** **DNA Gel Apparatus Bio-Rad (Wide-Mini-Sub® Cell GT)**	The Leinster Technology Centre, Blessington Industrial Estate, Blessington, Co. Wicklow, Ireland.
Heraeus Hera-Safe Laminair flow cabinet **Thermo Forma 3110 Water Jacketed CO_2 incubator**	Thermo Scientific, 12-16 Sedgeway Business Park, Witchford, Cambridgeshire CB62HY, England.
Hermle Z233M-2 Air-cooled version Microcentrifuge	Hermle Labortechnik GmbH, 25 Siemensstrasse, Wehingen, 78564, Germany.
Static Incubator	Sanyo Europe Ltd, 18 Colonial Way, Watford WD244PT, England.
Homogeniser Ultra Turrax	Janke and Kunkel IKA-Werk Ultra-Turrax, Staufen, 79129, Germany.

New Brunswick Scientific-Excella® Incubator Shaker Type E24 **Lauda AQUAline AL Water Bath** **New Brunswick Scientific Excella® E25 Orbital Incubator Shaker** **ND-1000 Spectrophotometer (Nanodrop)** **Eppendorf 5810r Benchtop centrifuge**	Mason Technologies, Greenville Hall, 228 South Circular Road, Dublin 8, Ireland.
Biacore 3000™ Instrument **Biacore sensor CM5 chips**	GE Healthcare Bio-Sciences AB, Danmarksgatan 41, M5, 75184 Uppsala, Sweden.
UNiPex Protein Arrays	Source BioScience GmbH, Campus Berlin-Buch, Robert-Rössle-Str.10, Erwin-Negelein-Haus (Building 79), 3125 Berlin, Germany.
Thermostatic Water Bath Model Y6	Grant Instruments (Cambridge) Ltd, 29 Station Road, Shepreth,

	Royston, Herts., SG8 6PZ, England.
Sigma 2K15 Benchtop Centrifuge	Sigma Laborzentrifugen GmbH, An der Unteren Söse 50, 37520 Osterode am Harz, Germany.
VAPOUR-Line ^{eco}50 (Vertical Autoclave) **Biometra $T_{GRADIENT}$ Thermal Cycler** **Millipore Ultra Pure H_2O Milli-Q Academic**	VWR International, Orion Business Campus, Northwest Business Park, Ballycoolin, Blanchardstown, Dublin 15, Ireland.
IKA® MTS 2/4 Digital Microtitre Shaker **Stuart Roller Mixer – STR1** **Stuart Mini See Saw Rocker SSM4** **Stuart Platform Shaker STR6** **Distinction Water Still D4000**	Lennox, John F. Kennedy Drive, Naas Road, Dublin 12, Ireland.

Tecan Safire^2 Plate reader	Tecan Austria GmbH, Untersbergstrasse, 5082 Grödig, Austria.

2.1.2 Reagents, kits and antibodies

All reagents were of analytical grade and purchased from Sigma-Aldrich Ireland Ltd. (Vale Road, Arklow, Wicklow Ireland), unless otherwise stated. All plastics were purchased from Sarstedt Ltd. (Sinnottstown Lane, Drinagh,Co Wexford, Ireland) unless otherwise stated.

Reagent	Supplier
Acetic acid **Agarose** **Dulbecco's Phosphate Buffered Saline (without Magnesium or Calcium)** **Presept tablets** **Carbenicillin** **Sodium Chloride**	Fisher Scientific Ireland, Suite 3 Plaza 212, Blanchardstown Corporate Park 2, Ballycoolin, Dublin 15, Ireland.

Hydrochloric acid **Macherey-Nagel Nucleospin Plasmid midiprep kit** **National Diagnostics™ Acrylagel™ 30% Solution (w/v)** **National Diagnostics™ Bis Acrylagel™ 2% Solution (w/v)** **Nucleospin Gel and PCR Clean up Kit**	
Bacteriological Agar **Tryptone Powder** **Yeast Extract**	Cruinn Diagnostics Ltd., Hume Centre, Parkwest Business Park, Nangor Road, Dublin 12, Ireland.
Trizol **Heat-Inactivated Foetal Bovine Serum** **S.O.C. Medium** **Ambion RNA*later*** **Superscript III cDNA synthesis kit** **M13KO7 Helper Phage**	Bio-sciences, Crofton Road, Dun Laoghaire, Dublin, Ireland.
Virkon **SureBlue Reserve TMB Microwell Peroxidase Substrate**	Lennox Laboratory Supplies Ltd., John F. Kennedy Drive, Naas Road, Dublin 12,

	Ireland.
Genomic DNA Transport Buffer	LGC Standards, Queens Road, Teddington, Middlesex, TW11 0LY, United Kingdom.
Milk Marvel	Premier Foods, Premier House, Centrium House, Griffiths Way, St Albans AL1 2RE, United Kingdom.
PCR Primers	Integrated DNA Technologies, Interleuvenlaan 12A B-3001, Leuven, Belgium.
New England Biolabs SfiI Restriction Enzyme and Buffer **New England Biolabs DNA ligase and Buffer** **New England Biolabs Antarctic Phosphatase and Buffer**	Brennan and Co., 61 Birch Avenue, Stillorgan Industrial Park, Stillorgan, Dublin, Ireland.
Complete™ Ultra Protease Inhibitor Tablets	Roche Diagnostics GmbH,

	Sandoferstrasse 116, DE-68305 Mannheim, Germany.
Attophos alkaline phosphatase substrate system **Amintra Ni^{+}NTA agarose resin**	MyBio Ltd. Kilkenny Research and Innovation Centre, St. Kieran's College Road, Kilkenny, Ireland.
MyTaq DNA Polymerase Red **MyTaq Red Mix Buffer** **Bioline DNA Hyperladder 1kb** **5x DNA Loading Buffer**	Medical Supply Company Ltd., Damastown, Mulhuddart, Dublin 15, Ireland.

Antibody	Supplier
Rabbit Anti-M13 + Fd Bacteriophage Coat Proteins pAb (Product - ab6188)	Abcam, 330 Cambridge Science Park, Cambridge, CB4 0F, England.

Mouse Anti-PolyHistidine Horseradish Peroxidase-conjugated mAb (Product - A7058) **Rabbit Anti-HA-Tag pAb (Product - SAB4300603)** **Goat Anti-Rabbit IgG Horseradish Peroxidase-conjugated pAb (Product – A0545)** **Rabbit Anti-Chicken IgY Horseradish Peroxidase-conjugated pAb (Product - A9046)**	Sigma Aldrich Ltd., Vale Road, Arklow, Co. Wicklow, Ireland.
Rat Anti-HA Horseradish Peroxidase-conjugated mAb (Product - 12013819001)	Roche Diagnostics Corporation, 330 Cambridge Science Park, Cambridge, CB4 0FL, England.
Mouse Anti-M13 Horse Radish Peroxidase-conjugated mAb (Product - 45-001-419) **Rabbit anti-Chicken IgY pAb (Product - 31104)**	Fisher Scientific Ireland, Suite 3 Plaza 212, Blanchardstown Corporate Park 2, Ballycoolin, Dublin 15, Ireland.

2.1.3 Mammalian cell lines

Cell Line	Tissue Type	Origin	Supplier
22Rv1	Prostate Carcinoma	CWR22R line	Applied Biochemistry Group, Dublin City University, Dublin, Ireland.
LNCaP	Prostate Carcinoma	Lymph Node metastasis	Applied Biochemistry Group, Dublin City University, Dublin, Ireland.
LNCaP C4-2	Prostate Carcinoma	LNCaP line	Prof. Thalmann, Bern University Hospital, Bern, Switzerland.
PC3	Prostate Adenocarcinoma	Metastatic bone tissue	Prof. Watson, Conway Institute, University College Dublin, Dublin, Ireland.

2.1.5 Bacterial cell strains

Strain	Genotype	Supplier
***E. coli* XL1-Blue Electrocompetent Cells**	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)]*	Agilent Technologies, 5301 Stevens Creek Blvd, Santa Clara, CA 95051, United States
One Shot TOP10F´ Chemically Competent *E. coli*	*F´{lacIq Tn10 (TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG*	Bio-sciences, Crofton Road, Dun Laoghaire, Dublin, Ireland.
***E. coli* SCS1 Supercompetent Cells**	*recA1 endA1 gyrA96 thi-1 hsdR17 (rK – mK +) supE44 relA1*	Agilent Technologies, 5301 Stevens Creek Blvd, Santa Clara, CA 95051, United States.

2.1.6 Commercial recombinant human protein expressing bacterial cells

All commercial recombinant human protein expressing clones were purchased from Source Bioscience (Source BioScience GmbH, Campus Berlin-Buch, Robert-Rössle-Str.10, Erwin-Negelein-Haus (Building 79), 3125 Berlin, Germany) encoded on pQE30NST plasmids in SCS1 cells.

Protein	Clone ID	Genbank ID
Proteasome Subunit Alpha Type-7	IMGSp9027C1115D	AK127210
SET Domain Containing 2	IMGSp9027C0934D	BX649110
Zinc Finger 358	IMGSp9027B1237D	NM_018083
Retinoic Acid Receptor Gamma	IMGSp9027C0873D	NM_000966
RAS dexamethasone induced	IMGSp9027A0931D	BC042688

2.1.7 Media and buffers

2.1.7.1 General buffers

Buffer	Composition
Phosphate-buffered saline (PBS)	0.15 M NaCl 2.5 mM KCl

	10 mM Na_2HPO_4 18 mM KH_2PO_4 Constituents are dissolved in 800 mL ultra-pure H_2O, the pH is adjusted to 7.4, by titration with appropriate acid or base and finally made up to 1 litre with ultra-pure H_2O.
PBS-Tween 20 (0.05%) (PBST)	0.15 M NaCl 2.5 mM KCl 10 mM Na_2HPO_4 18 mM KH_2PO_4 0.05% (v/v) Tween 20 pH 7.4
FACs Buffer	PBS formulation above 20mM HEPES 1% (w/v) BSA 0.03% (w/v) NaN_3 pH 7.4 0.2μm Filter Sterilised
Tris-buffered saline (TBS)	50 mM Tris-HCl 150 nM M NaCl Constituents are dissolved in 800 mL ultra-pure H_2O, the pH is adjusted to 7.5, by titration with appropriate acid or base

	and finally made up to 1 litre with ultra-pure H_2O.
TBS-Tween 20 (0.05%) (TBST)	50 mM Tris-HCl 150 nM M NaCl 0.05% (v/v) Tween 20 pH 7.5
TBST-TritonX-100	50 mM Tris-HCl 150 nM M NaCl 0.05% (v/v) Tween 20 0.5% (v/v) TritonX-100 pH 7.5
AP Buffer	0.1 M Tris Base 1 mM $MgCl_2$ Constituents are dissolved in 800 mL ultra-pure H_2O, the pH is adjusted to 7.5, by titration with appropriate acid or base and finally made up to 1 litre with ultra-pure H_2O.
HBS-EP$^+$	0.1 M HEPES 1.5 M NaCl 30 mM EDTA 0.05% (v/v) Tween 20 Constituents are dissolved in MilliQ ultra-pure H_2O, the pH is adjusted to 7.4, by

	titration with appropriate acid or base and sterile filtered (0.2 μm filter) and degassed prior to use.
(Milk/BSA)PBST/TBST	PBST/TBST formulation above Specified % (w/v) BSA/Milk Marvel™ powder pH 7.4

2.1.7.2 Bacterial culture formulations and supplements

Media	Formulation
Luria Bertani Broth (LB) medium	Tryptone 10 g/L Yeast Extract 5 g/L NaCl 5 g/L
Super Broth (SB) medium	MOPS (pH 2.5-4) 10 g/L Tryptone 30 g/L Yeast Extract 20 g/L
ZY medium	Tryptone 10 g/L Yeast Extract 5 g/L
1 x 505 media supplement	0.5% (v/v) Glycerol 0.05% (v/v) Glucose
50 x 5052 media supplement	Glycerol 250 g/L Glucose 25 g/L α-Lactose 100 g/L
20 x NPS	$(NH_4)_2SO_4$ 66 g/L

	KH_2PO_4 136 g/L Na_2HPO_4 142 g/L
Solid medium was made by adding 15 g/L bacteriological agar, or 7.5 g/L for soft agar, to the media formulations above.	

2.1.7.3 Reagent preparation for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Preparation of 12.5% Separation Gel	1 gel (6 mL)
1 M TrisHCl (pH 8.8)	1.5 mL
30% (w/v) Acrylagel	2.5 mL
2% (w/v) Bis Acrylagel	1.0 mL
dH_2O	934 µL
10% (w/v) sodium dodecyl sulfate (SDS)	30 µL
10% (w/v) Ammonium persulfate (APS) (Fresh)	30 µL
N,N,N',N'-Tetramethylethylenediamine (TEMED)	6 µL

Preparation of 4.5% Stacking Gel	1 gel (2.6 mL)
1 M Tris HCl (pH 6.8)	300 µL
30% (w/v) Acrylagel	375 µL
2% (w/v) Bis Acrylagel	150 µL
dH_2O	1.74 mL
10% (w/v) SDS	24 µL
10% (w/v) APS (Fresh)	24 µL

TEMED	2.5 μL

10X Electrophoresis Buffer	1L
196 mM Glycine	144 g
50 mM Tris Cl (pH 8.3)	30 g
0.1% (w/v) SDS	10 g

4X Protein Loading Dye	10 mL
Tris 0.5 M (pH 6.8)	2.5 mL
Glycerol	2 mL
2-mercaptoethanol	0.5 mL
20% (w/v) SDS	2.5 mL
Bromophenol Blue	20 ppm (1.9 mg)
dH_2O	2.5 mL

2.1.7.4 Reagent preparation for immobilised metal affinity chromatography

Lysis Buffer A	100 mL
300 mM NaCl	1.754 g
50 mM NaH_2PO_4	0.69 g
10 mM Imidazole	0.068 g
Adjust pH to 7.5, using 1 M NaOH	

Lysis Buffer B	100 mL
100 mM NaH_2PO_4	1.38 g

10 mM TrisCl	0.12 g
6 M Guanidine hydrochloride (GuHCl)	57.3 g
Adjust pH to 8, using 1 M NaOH	

Wash Buffer A	100 mL
50 mM NaH_2PO_4	0.69 g
1 M NaCl	5.844 g
10% (v/v) glycerol	20 mL 50% (v/v) sol.
20 mM Imidazole	0.136 g
1% (v/v) Triton X-100	1 mL
Adjust pH to 7.5, using 1 M NaOH	

Wash Buffer B	100 mL
50 mM NaH_2PO_4	0.69 g
300 mM NaCl	1.754 g
20 mM Imidazole	0.136 g
Adjust pH to 7.5, using 1 M NaOH	

Wash Buffer C	100 mL
100 mM NaH_2PO_4	1.38 g
10 mM TrisCl	0.12 g
8 M Urea	40.05 g
Adjust pH to 6.3, using 1 M HCl	

Elution Buffer A	100 mL
50 mM NaH_2PO_4	0.69 g
300 mM NaCl	1.754 g
250 mM Imidazole	1.7 g
Adjust pH to 7.5, using 1 M NaOH	

Elution Buffer B	50 mL
100 mM NaH_2PO_4	0.69 g
10 mM TrisCl	1.754 g
8M Urea	1.7 g
Adjust pH to 4.5, using 1 M HCl	

2.1.8 Plasmids for cloning

The pComb3XSS vector was made available to the Applied Biochemistry Group by Professor Carlos Barbas III of the Scripps Institute, La Jolla, California, USA. The vector contains a 1.6 kb stuffer fragment that can be removed with an SfiI restriction enzyme, allowing placement of scFv DNA into the vector. To ensure the vector was still intact, it was treated with SfiI and run on a 1% agarose gel prior to use. The vector is depicted in Figure 2.1 below.

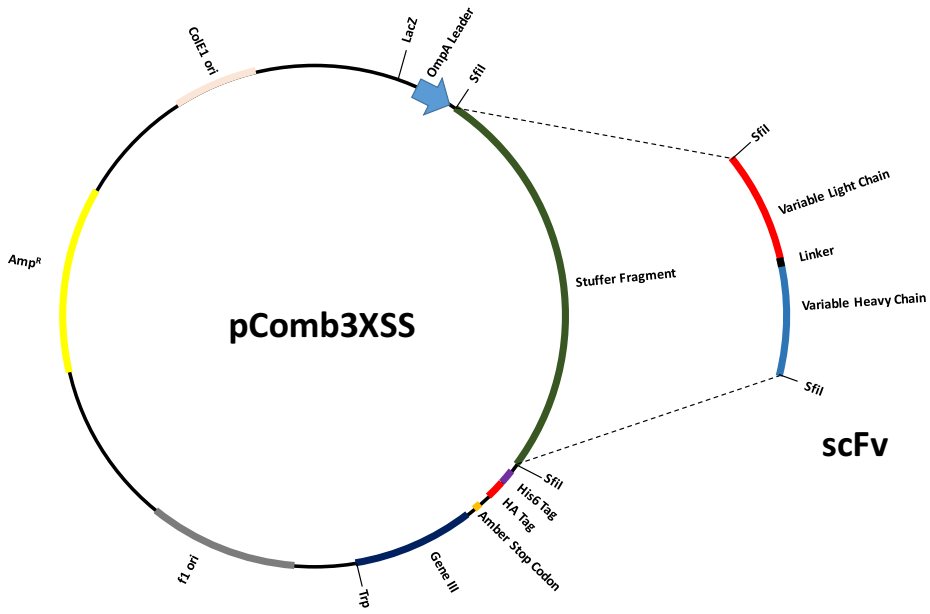

Figure 2.1 - Schematic of the pComb3XSS vector. The vector is ~5kb in size and contains a ~1.6kb Stuffer fragment by two SfiI sites. The scFv gene replaces the stuffer fragment and precedes the gene III to allow for scFv fusion to the pIII coat protein on phage. An Amber Stop Codon lies between the stuffer fragment and gene III to allow for soluble expression in non-suppressor strains. Before the Amber Stop Codon are HIS and HA tags to allow for scFv purification with commercially available kits. Other traits of interest granted by the plasmid are ampicillin resistance, a LacZ promoter and both bacterial and phage origins of replication.

2.2 Methods

2.2.1 Cell culture

All cell culture work was performed in Class II Laminar Flow Cabinets using the required level of Personal Protective Equipment (Disposable Gloves, Lab Coats and Protective Eye Glasses). Additional Personal Protective Equipment was used when working with liquid nitrogen (Extreme temperature resistant gloves and face guard).

2.2.2 Cell lines used

The cell lines used in this research were the androgen-responsive lines LNCaP and 22Rv1 and the androgen-unresponsive lines PC3 and LNCaP C4-2. The LNCaP line was originally isolated from metastatic prostate carcinoma of a patient's lymph node. The LNCaP C4-2 line was derived from a tumour that developed in a castrated nude mouse injected with LNCaP cells (Thalmann *et al.*, 1994). The line 22Rv1 is a prostate carcinoma-derived from an androgen-dependent CWR22 xenograft. The PC-3 line was established from a grade 4 prostatic adenocarcinoma. Both the 22Rv1 and LNCaP cell lines were taken from the cryopreserved stocks of the Applied Biochemistry Group in Dublin City University, Ireland. PC-3 cells were kindly provided by Professor William Watson, University College Dublin, Ireland. LNCaP C4-2 cells were a gift from Professor George Thalmann, Bern University, Switzerland. All cell lines were authenticated and validated as 'mycoplasma-free' prior to use.

All cell lines are adherent and were grown in vented T-25, T-75 and T-175 flasks (Sarstedt) in 5% (v/v) CO_2 at 37^{o}C. All manipulations were carried out aseptically in Laminar Flow cabinets. All cell lines were cultured for a minimum of 3 passages before

being used in any experimental work. Only cells with a passage number lower than 40 were used in this work.

2.2.3 Medium preparation

All lines were maintained in RPMI 1640 with L-glutamine (0.3 g/L) and $NaHCO_3$ (2 g/L) supplemented with 10% (v/v) heat inactivated Foetal Bovine Serum (FBS), penicillin (100 IU/mL)/streptomycin (10 mg/mL) and 1mM Sodium pyruvate. All media were stored at 4^{o}C. Upon delivery, supplements were divided into single-use aliquots and were stored at -20^{o}C.

2.2.4 Culture of cell lines

All cell lines were grown in supplemented RPMI media (see above) to ~70% confluency. Trypsinisation was required prior to sub-culture, as all cell lines used are adherent. All reagents added to cells were pre-warmed to 37^{o}C in a water bath.

During sub-culturing, the medium was removed from the T-flask and the surface was washed with Dulbecco's Phosphate Buffered Saline ('Magnesium and Calcium-free') (DPBS) and incubated for 5-10 minutes at 37^{o}C in the presence of trypsin-EDTA (0.5 g/l porcine trypsin and 0.2 g/l EDTA-4Na in Hank's Balanced Salt Solution with phenol red). The volume of trypsin added varied from 2-5mL depending on the size of the T-flask (1mL for T-25, 2mL for T-75 and 5mL for T-175). Upon confirmation of cell detachment using a 40X inverted bright field microscope, the trypsin-EDTA was then neutralised using an equal volume of FBS-containing media and the cells were pelleted by centrifugation for 5 minutes at 500 g (Sigma 2K15 centrifuge).

The pellet was resuspended in fresh medium and 20-25% (determined after the confluence was investigated) of the population was seeded in a new T-flask with fresh medium. In the case of up-scaling a culture into a T-flask of a larger volume, 100% of the cell population was seeded. Three days were typically required before a cell line required subculture, although this time period-varied.

2.2.5 Cryopreservation and recovery of cell lines

For long-term storage of cell lines, a solution of standard medium was used (see Section 2.2.3), apart from the FBS content being increased from 10% (v/v) to 50% (v/v) in cryopreservation media. For cryopreservation, this solution was supplemented with 10% (v/v) dimethyl sulfoxide (DMSO).

Cells were harvested from the surface of a T-175 using 5mL of trypsin. Following trypsin neutralization, the cells were spun down by centrifugation for 5 minutes at 500 g (Sigma 2K15 centrifuge). The pellet was resuspended in previously prepared cryopreservation solution and dispensed into 2mL cryovials. Typically, each vial would be used to store 1mL of freeze medium containing 5-10 x 10^6 cells. Vials were placed in a Nalgene® Mr. Frosty and stored for 12 hours at -80^{o}C for slow freezing before being transferred to liquid nitrogen for long-term storage.

For recovery of cell lines from storage in liquid nitrogen, the relevant cryovial was placed in a 37^{o}C water bath for fast thawing. Then, the total 1 mL aliquot of cryopreserved cells in freeze medium was added to 4 mL of prewarmed-medium to dilute the DMSO. The cells were pelleted by centrifugation for 5 min at 500 g in a Sigma 2K15 benchtop

centrifuge. The supernatant was removed and the pellet was resuspended in 5mL of supplemented medium (FBS concentration increased to 50% (v/v)) warmed to 37^{o}C. The solution was placed in a T25 flask and left overnight in 5% (v/v) CO_2, and at 37^{o}C to adhere. The following day, the supernatant was replaced with fresh prewarmed medium. After a 24 to 36-hour period, the cells were examined under a Nikon Diaphot Inverted Tissue Culture Microscope to confirm they were adherent and growing well. Following confirmation, the FBS content of the media was reduced from 50% (v/v) to 10% (v/v) and the cells were cultured using the standard protocol (see Section 2.2.4).

2.2.6 Cell counting and viability

Cells were counted using a Neubauer C-Chip Haemocytometer. Suspended cells (50μL) were diluted 1 in 2 with a solution of Trypan blue in water (0.25% (w/v). Immediately after Trypan blue staining the solution was loaded into the haemocytometer and analysed at a 20-40X magnification using a light microscope. This process would typically take 10-15 min. Cells stained blue were considered dead whereas white cells were considered viable.

Viable cells in each of the 5 large squares in the haemocytometer chamber were counted and an average obtained. This number was multiplied by the dilution factor (2) and the volume factor (10^4) to determine the number of cells/mL.

2.2.7 Mycoplasma testing

All lines were tested for mycoplasma using the MycoAlert PLUS Mycoplasma Detection Kit (Lonza, LT07-705) in accordance with the protocol developed by the manufacturer.

The assay works by lysing any mycoplasma present in the culture supernatant and releasing mycoplasmal enzymes. These enzymes, if present, catalyse the conversion of adenosine diphosphate (ADP), present in the kit substrate, to adenosine triphosphate (ATP). Luciferase was then added to convert the ATP present into a light signal that is quantified using a luminometer.

2.2.8 Cell line authentication

Genomic DNA from the cell lines PC3, 22Rv1, LNCaP and LNCaP C4-2 was sent to LGC Standards (Middlesex, UK) for authentication using the PowerPlex® 16 Loci Service prior to experimental use. This service measures the length of the following 15 Short Tandem Repeats (STR); Penta E, D18S51, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, vWA, Penta D, CSF1PO, D16S539, D7S820, D13S317 and D5S818 and Amelogenin as a gender determinant. The STRs and amelogenin from the genomic DNA were profiled and compared against a database of 3000 cell line STR profiles. Genomic DNA was sent to the service provider by preparing 2 x 10^6 cells from live cultures according to the service providers specifications. Each cell line was pelleted through centrifugation at 500 g for 5 min in a Sigma 2K15 Benchtop Centrifuge and the pellet was resuspended in 400 µL of manufacturer provided transport buffer. The transport buffer lyses the cells and preserves the genomic DNA during the transport of the materials to the service provider. Reference profiles were available for PC3, 22Rv1 and LNCaP lines, to which the genomic DNA matched. There was not, however, a database reference available for the LNCaP C4-2 line and the STR profile did not match any of the 3000 profiles in the database. As LNCaP C4-2 is a sub-line of LNCaP, it was compared against the LNCaP STR profile to which there was a significant overlap, as expected.

2.2.9 Lysis of cell lines

To prepare cell lysate, the cell line of interest was grown to ~90% confluency in two T-175 culture flasks. The medium was removed and the surface containing the adhered cells was washed with DPBS. Cell Dissociation Solution (Non-enzymatic) (Sigma) was added to the surface of the T-flask and allowed to incubate at 37^{o}C, 5% (v/v) CO_2. Following confirmation of cell detachment via light microscopy, the supernatant was removed. Any remaining cells were removed using a cell scraper in the presence of DPBS. The cells were pelleted by centrifugation for 5 min at 250 g in a Sigma 2K15 benchtop centrifuge and resuspended in 5 mL fresh DPBS. This centrifugation and resuspension step was repeated. The cells were then pelleted as before and resuspended in 1 mL of ice cold RIPA (Radio-immunoprecipitation assay) buffer (Sigma, R0278) with a 1 in 100 dilution of mammalian cell protease inhibitor cocktail (Sigma, I3786). Samples were homogenised by pelleting through centrifugation at 10,000 g in a Hermle benchtop centrifuge and resuspended, without removing the buffer, by repeatedly flushing the sample with a 25G needle and syringe. This step was repeated twice. The supernatant was then sonicated on ice for 3 seconds at 15% amplitude with 25 second intervals for a total of 5 min, using a Branson Digital Sonifier 250 microtip-sonicator. The protein concentration of the resulting supernatant was determined by measuring the absorbance at 280 nm with a Nanodrop 1000 and the supernatant was aliquoted into multiple 200 μL volumes and stored at -80^{o}C.

2.2.10 Cell preparation for chicken immunisation and phage panning

Trypsin is a proteolytic enzyme used to dislodge adherent cells by cleaving anchoring proteins. Hence, for assays that require intact surface markers, an enzyme-free

dissociation solution (Sigma) was used to detach cells. The dissociation solution uses a mixture of chelators to gently dislodge adherent cells from their growth surface. To achieve cell detachment without the use of trypsin, the culture supernatant was removed from the T-flask and the surface washed with DPBS. Dissociation solution was then added to the flask and left to incubate in 5% (v/v) CO_2 at 37^{o}C. The flask was checked periodically for cell detachment from the surface using an inverted microscope. Typically, the process would take 30-45 minutes before satisfactory detachment. The cells were then pelleted by centrifugation for 5min at 500 g in a Sigma 2K15 benchtop centrifuge. The dissociation solution was removed and the pellet resuspended in PBS. A cell count and a viability assay were performed and the required number of cells was taken to mix with adjuvant prior to immunisation (see Table 2.1).

2.3 Generation of an avian scFv library against PCa cell lines

2.3.1 Research ethics

All animal procedures were ethically approved by the Dublin City University Research Ethics Committee (REC number - DCUREC/2011/016) and carried out in accordance with the Department of Health and Children/Cruelty to Animals Act Licence number B100/2705. Procedures were carried out under the direction of trained animal handlers whilst minimising animal distress. A Laboratory Animal Science and Training program (LAST) was completed prior to the commencement of any animal work. All animals were handled and treated humanely in the dedicated Bio-Resource Unit (BRU) in Dublin City University.

2.3.2 Immunisation of two leghorn chickens with PCa cells

Seven days prior to immunisations, serum was taken from both chickens (ABG055 and ABG048) as a pre-immunisation bleed. Additionally, serum was extracted from both chickens 5-7 days following immunisations procedures, to compare IgY titres. For preparation of serum samples, bloods taken from the chickens were left at room temperature for 2 hours to clot and stored overnight at 4^{o}C. The following morning the clots were removed using sterile tweezers and the samples were centrifuged at 2,057 g in an Eppendorf 5810r Benchtop centrifuge for 30 min. The supernatant was aliquoted and stored at -20^{o}C for short-term storage and -80^{o}C for long-term storage.

Differing PCa cell line immunisation strategies were applied to two Leghorn chickens (ABG048 and ABG055). The strategy employed for ABG048 (immunised with all four available PCa cell lines) was intended to generate an antibody library against proteins common to all PCa cell lines. The strategy employed for ABG055 (PC3 and LNCaP C4-2) was focused on androgen-independent forms of the disease (see Table2.1). Chickens were immunised over a 3-month period, with boosts 6 and 13 weeks after the primary immunisation. Bleeds were taken prior to the initial immunisation. Bleeds were also taken 1 week after each boost. Following confirmation of an immune response against PCa cells, a final boost was administered less than 1 week before sacrifice and RNA extraction.

Table 2.1 - Chicken immunisation schedule.

Chicken	Week	Procedure	Cell Lines used	Number
ABG048	1	Primary Immunisation	LNCaP, 22Rv1, LNCaP C4-2, PC3	~1 x 10^7, ~1 x 10^7, ~1 x 10^7, ~1 x 10^7
	6	Boost 1	PC3	~1.5 x 10^7
			LNCaP C4-2	~1.5 x 10^7
	13	Boost 2	LNCaP	~1.5 x 10^7
			22Rv1	~1.5 x 10^7
	23	Boost for sacrifice	LNCaP, 22Rv1, LNCaP C4-2, PC3	~1 x 10^7, ~1 x 10^7, ~1 x 10^7, ~1 x 10^7
ABG055	1	Primary Immunisation	PC3	~2 x 10^7
			LNCaP C4-2	~1 x 10^7
	6	Boost 1	PC3	~1.5 x 10^7
			LNCaP C4-2	~1.5 x 10^7
	13	Boost 2	PC3	~1.5 x 10^7
			LNCaP C4-2	~1.5 x 10^7
	23	Boost for sacrifice	PC3	~2 x 10^7
			LNCaP C4-2	~2 x 10^7

For immunisations, PCa cells were prepared suspended in 500μL of DPBS (see Section 2.2.10). The cell suspension was mixed 1:1 with Freud's Complete Adjuvant for the primary immunisation (Sigma), and with Freud's Incomplete Adjuvant (Sigma) for booster immunisations. Emulsification was achieved by passing the mixture through a needle on a syringe, followed by vortexing, until a consistent milky suspension was obtained. The immunogen was administered to the chickens subcutaneously over seven

different sites. Following an appropriate immune response, determined by a serum titre ELISA, the chickens were given a final boost one week prior to sacrifice and RNA extraction.

Chickens were sacrificed using intravenous administration of Pentobarbital Sodium (Euthathal). Surgical tools were prepared for RNA extraction, before sacrifice, by incubation in 1X Presept™ for 30 minutes, rinsing in Molecular Grade Water (Sigma) followed by being autoclaved at 121^{o}C for 20 min. Finally, the tools were incubated overnight on a rollershaker, soaking in RNAse Zap. RNAse Zap removes contaminating RNAses from surfaces, which may lead to degradation of RNA. The chickens' spleens were then removed with the surgical tools (scissors, scalpels and tweezers). The spleens were then placed in 'RNase-free' universal tubes and submerged in RNA*later*® for 30 min at -20^{o}C to allow tissue penetration. The tubes were then flash frozen in liquid nitrogen and stored at -80^{o}C until RNA extraction was to be performed. RNA*later*® permeates tissues and cells to stabilize and protect internal RNA. The solution remains liquid at -20^{o}C, allowing for tissue penetration before flash freezing. At temperatures of -20^{o}C or lower, RNA can be stored indefinitely.

2.3.3 Avian antibody serum titre against PC3 cells

For antibody titre assays PC3 cells were seeded 100,000 cells per well into a Nunc 96, clear well, cell culture plate in 200μL medium (see Sections 2.2.4 and 2.2.6, for detachment and counting protocol). The cells were then left to incubate at 37^{o}C in 5% (v/v) CO_2 overnight to adhere. The following day the, the medium was removed and the wells were gently washed with 100 μL DPBS taking care not to remove any adhered cells.

Following confirmation of surface confluence using light microscopy, 100μL of 8% (v/v) formalin in DPBS was added to each well for fixation for 25 min at 37°C. The plate was washed three times with dH_2O, then three times with Phosphate Buffered Saline (PBS). The plate was blocked with 3% (w/v) Bovine Serum Albumin (BSA) in PBS for 1 hour at 37°C and washed three times with PBS. Serial dilutions (up to 1 in 1,000,000) of avian IgY-containing sera in 1% (w/v) BSA PBS Tween 0.05% (v/v) (PBST) were added to designated wells and left for 90 min at room temperature to bind. The wells were then washed three times with PBST and three times with PBS. A 1 in 2,000 dilution of Horse Radish Peroxidase (HRP)-labelled rabbit anti-chicken IgY (Sigma) in 1% (w/v) BSA PBST (100 μL) was added to each well and left at room temperature for 90 min. The plates were washed three times with PBST and three times with PBS. 3,3',5,5'-Tetramethylbenzidine (TMB) with H_20_2 substrate solution (100μL) was added to each well and left at room temperature until a visible colour change was detectable (~15min), and the reaction then stopped by addition of 50 μL of 10% (v/v) HCl. The absorbance was read at 450 nm using a Tecan Safire2 plate reader.

2.3.4 Isolation of total RNA from spleens of each chicken

Prior to the extraction procedure, the following preparative steps were taken. Oakridge tubes were cleaned with Virkon® to ensure that no potential contaminants were present. The tubes were then washed with 'RNase-free' Molecular Grade water (Sigma), sprayed with RNase zap® (Sigma) and left overnight to ensure no contaminating RNases remained. The homogenizer parts were incubated in 1X Presept™, washed thoroughly with 'Rnase-free' water, autoclaved (121°C for 20 min) and baked in an oven at 60°C overnight. Molecular grade ethanol (75%, v/v) in 'Rnase-free' water) was prepared in 'RNase-free' tubes and stored at -20°C. A Biological Safety Cabinet (Class II) was de-

contaminated by cleaning with 70% (v/v ethanol in 'RNase-free' water) Industrial Methylated Spirits and RNase Zap®. Materials and reagents to be used during the extraction procedure were also decontaminated. The spleens, previously stored at -80°C, were moved to -20°C overnight so the RNA later would be in a liquid state and easily removed during the procedure. All manipulations were performed aseptically in a Laminar flow cabinet. All reagents were chilled at either -20°C or on ice before being used. All tubes and equipment treated with RNase Zap prior to the procedure were rinsed with 'RNase-free' molecular grade water just before use.

The following steps were carried out on the spleens of chickens ABG048 and ABG055. RNA*later*® was removed from the spleens by centrifugation at 1,600 g at 4°C for ten minutes in an Eppendorf 5810 benchtop centrifuge followed by removal of the supernatant. The spleens were placed in 'RNAse-free' tubes and homogenised in 30 mL of Trizol reagent using a sterile homogenizer and left for 5 min to allow for nucleoprotein complexes to dissociate. The tubes were centrifuged at 1,575 g for 10 min in an Eppendorf 5810 benchtop centrifuge at 4°C. The supernatant was removed into sterile 50 mL Oakridge tubes containing, 6 mL of molecular grade chloroform (Sigma), mixed well by shaking and incubated at room temperature for 15min. The samples were then centrifuged at 17,500 g for 25 min (Eppendorf 5810 benchtop centrifuge) at 4°C to produce three discernible layers. The RNA-containing upper aqueous layer was carefully removed, taking care not to transfer any proteins from the interphase. This was added to an Oakridge tube containing 15 mL of ice-cold isopropanol to precipitate the RNA from the solution. This solution was mixed and left at room temperature for 15 min. The tube was then centrifuged at 17,500g in an Eppendorf 5810 benchtop centrifuge for 25 min at 4°C to pellet the precipitated RNA. The supernatant was carefully removed and

discarded ensuring that the pellet was not disrupted. The RNA pellet was washed with the addition of 30 mL 75% (v/v) ethanol and was centrifuged again at 17,500 g for 10 min in an Eppendorf 5810 benchtop centrifuge. The supernatant was discarded and the pellet was allowed to air-dry in the laminar flow cabinet. Following air-drying, the pellet was gently resuspended in 500 μL of Molecular grade water. The RNA was quantified using a Nanodrop™spectrophotometer (based on measuring absorbance at 260 nm), and run for 45 minutes on a 1.5% (w/v) agarose gel (w/v) in Tris Acetate Ethylenediaminetetraacetic (EDTA), pH 8.3, for integrity analysis. The extracted RNA was then carried forward for reverse transcription cDNA synthesis (see Section 2.3.5).

2.3.5 cDNA synthesis via reverse transcription

The coding DNA (cDNA) required for amplification of variable light and heavy chain gene fragments was generated through reverse transcription polymerase chain reaction (RT-PCR) of RNA extracted from chicken spleens using the Superscript III first strand DNA synthesis kit (Invitrogen).

For each reaction, a total of 5μg of RNA was used. Extracted RNA from each spleen was used for 20 RT-PCR reactions. Two 20X mixtures were prepared for RT-PCR:

Mix 1	**1 reaction**	**20 reactions**
RNA	XμL (should contain 5μg of RNA)	20X
Oligo dTs	1μL	20μL
dNTPs (10mM)	1μL	20μL
Molecular Grade Water	To 10μL	To 200μL

Mix 2	**1 reaction**	**20 reactions**
10X RT Buffer	2μL	40μL
$MgCl_2$	4μL	80μL
0.1M DTT	2μL	40μL
RNase OUT (40U/μL)	1μL	20μL
Superscript III RT (200U/μL)	1μL	20μL

Mix 1 was split into eight 25 μL aliquots and incubated at 65^{o}C for 5 min (Biometra TGradient PCR machine). These were then kept on ice for 1 min. An equal volume of Mix 2 was added to each 25 μL sample of Mix 1 and incubated at 50^{o}C for 5 min. The reaction was continued at 85^{o}C for 5 min. RNase H (20 μL) was then added to each mixture for template RNA removal and the tubes were subsequently incubated at 37^{o}C for 20 minutes. Finally, the eight separate aliquots were combined and stored at -80^{o}C. The amount of cDNA present was then determined by quantification on a Nanodrop^{TM}Spectrophotometer.

2.3.6 Amplification of chicken antibody variable domain sequences using PCR

Primers, for PCR-based amplification of the variable light and heavy gene segments of ABG048 and AGB055 cDNA, described by Barbas for the pComb Vector system (Barbas *et al.,* 2001) were used. A linker was included in the CSCVHo-FL primer so that the C-terminus of the light chain can be linked to the N-terminus of the heavy chain, using a Splice by Overlap Extension (SOE). The longer linker fragment $(G_4S)_4$ is thought to improve antibody affinity and reduce aggregate formation compared to the shorter fragment (GGSSRSS), which tends to dimerise. PCR Primers were obtained 'lab-ready' from Integrated DNA Technology.

V_H Primers

CSCVHo-FL (sense), Long Linker

5' – GGT CAG TCC TCT AGA TCT TCC GGC GGT GGT GGC AGC TCC GGT GGT GGC GGT TCC GCC GTG ACG TTG GAC GAG -3'

CSCG-B (reverse)

5' – CTG GCC GGC CTG GCC ACT AGT GGA GGA GAC GAT GAC TTC GGT CC – 3'

V_L Primers

CSCVK (sense)

5' – GTG GCC CAG GCG GCC CTG ACT CAG CCG TCC TCG GTG TC – 3'

CKJo-B (reverse)

5' – GGA AGA TCT AGA GGA CTG ACC TAG GAC GGT CAG G – 3'

Overlap Extension Primers

CSC-F (sense)

5’ – GAG GAG GAG GAG GAG GAG GTG GCC CAG GCG GCC CTG ACT CAG – 3’

CSC-B (reverse)

5’ – GAG GAG GAG GAG GAG GAG GAG CTG GCC GGC CTG GCC ACT AGT GGA GG – 3’

Amplification of variable genes were performed using MyTaq Red™ polymerase (Bioline) and its corresponding Red Mix buffer. The enzyme is a high-yield polymerase that uses a pre-optimised buffer designed for fast high-throughput amplification of sequences <1000bp in length. The length of DNA fragments used for scFv generation range from ~340-800bp. The robustness of the enzyme also makes it suitable for colony pick PCRs.

For amplification of Variable Heavy and Variable Light Segments, the following reaction mixture was prepared:

Component	**Concentration**
Template cDNA	1μL
Forward primer	50μM
Reverse Primer	50μM
MyTaq Red Polymerase (5U/μL)	1μL
5x MyTaq Red Buffer	10μL
Molecular Grade Water	To 50μL

The following PCR program was set up:

Step	Temperature	Time	Cycles
1			1
Initial Denaturation	95°C	5 min	
2			30
Denaturation	95°C	15 sec	
Annealing	56°C	15 sec	
Extension	72°C	10 sec	
3			1
Final Extension	72°C	10 min	

2.3.7 Purification of PCR products

Following PCR, the products were loaded directly onto a 1.5-2% agarose gel (varies depending on expected product size) containing SYBR SAFE® (see Section 2.3.14) and run at 95V for 45 min alongside a 1kb HyperLadder (Bioline). Products were then visualised using a Dark Reader transilluminator (Clare Chemical Research) and, following confirmation of product presence based on size, the DNA bands were excised from the gel using a sterile, sharp scalpel and underwent purification using a NucleoSpin® PCR Clean Up Kit (Macherey-Nagel).

Excised gels were incubated at 50°C with kit Buffer NTI (200µL of buffer for every 100mg of gel) until the gel had completely dissolved. The solution was then passed through a NucleoSpin® Gel and PCR Clean-up Column by centrifugation (700µL at a time at 11,000 x g for 30 sec in a Hermle Microcentrifuge) and the flow-through discarded. Buffer NT3 (700 µL) was then added to the column and the column was centrifuged at 11,000 g for 30 seconds and the flow-through discarded. This wash step was repeated once more. The column was then centrifuged as before for 2 minutes to dry the DNA-bound membrane. Buffer NE was added to the column (volume depending on the desired concentration of DNA) and the column was placed into a clean microfuge tube and left to incubate at room temperature for 2 minutes. The DNA was eluted into the microfuge tube by centrifugation for 1 min at 11,000 g in a Hermle Microcentrifuge. A Nanodrop™1000 instrument (260nm) was then used to evaluate the concentration of the DNA. Samples were aliquoted at the desired volume and stored at -20°C.

2.3.8 ScFv fragment generation through splice by overlap extension (SOE)

Variable Heavy and Light Chain segments generated in the previous PCR have an overlapping sequence (contained in primers used) that joins them via the incorporation of a glycine-serine linker $(G_4S)_4$. For the SOE-PCR, equal parts of V_H and V_L fragments are mixed and an overlap product of ~750 - 800bp is generated.

For the SOE-PCR, the following reaction mixture was prepared:

Component	**Concentration**
V_H product	100ng
V_λ product	100ng
Forward primer	50μM
Reverse Primer	50μM
MyTaq Red Polymerase (5U/μL)	1μL
5x MyTaq Red Buffer	10μL
Molecular Grade Water	To 50μL

The PCR program was set up:

Step	Temperature	Time	Cycles
1			1
Initial Denaturation	95°C	5 min	
2			25
Denaturation	95°C	15 sec	
Annealing	56°C	15 sec	
Extension	72°C	10 sec	
3			1
Final Extension	72°C	10 min	

Following the SOE-PCR, the products were loaded onto a 1.5% (w/v) agarose gel (see Section 2.3.14) and run under standard conditions for confirmation of SOE amplification. For large-scale amplification, products were run for a longer period on a larger gel to ensure adequate band isolation and to aid excision. The SOE product was purified and quantified using the same method as for the V_H and V_L fragments (see Section 2.3.6). For library generation, at least 10 μg of SOE was generated before restriction and ligation procedures, as DNA can be lost during steps preceding transformation.

2.3.9 Transformation and purification of pComb3XSS vector for cloning

Sufficient stocks of the vector pComb3XSS to be used for SOE insertion needed to be generated and purified prior to restriction procedures. For this, a stock of previously purified plasmid was obtained from the Applied Biochemistry Group's stocks and transformed into electro-competent *E. coli* XL1-Blue cells. For this, 10 ng of the vector

stock was transformed into 100 μL of cells using the same transformation protocol as in section 2.3.16. Following a 1-hour rescue in SOC medium at 37°C, serial dilutions were performed and 100 μL of each solution was plated onto LB agar Carbenicillin (100 μg/mL) plates to obtain single colonies.

Single colonies were picked and grown in 10 mL LB Carbenicillin (100 μg/mL) cultures, overnight, in a shaking incubator at 300 rpm and 37°C. The following day 100 μL from each culture was sub-cultured into 100 mL of SB Carbenicillin (100 μg/mL) and grown overnight at 37°C, while shaking at 300 rpm. The plasmids contained in each culture were then purified using a Macherey-Nagel NucleoBond Xtra Midi kit according to the manufacturer's instructions. Briefly, each culture was pelleted by centrifugation for 30 min at 3,330 x g in an Eppendorf 5810 benchtop centrifuge and the supernatant was removed. The pellet was resuspended in 8 mL RES buffer with RNase A. Buffer LYS (8 mL) was added to lyse the cells and the solutions were mixed by gentle inversion 5 times prior to incubating for 5 min at room temperature. The Buffer LYS was neutralised with the addition of Buffer NEU (8 mL) and mixed 15 times by gentle inversion. The solution was loaded onto a previously equilibrated NucleoBond® Xtra Column filter apparatus. The kit filter column was washed with Buffer EQU (5 mL) and the column discarded. The filter was washed with 8 mL Buffer WASH before eluting the filter-bound plasmid DNA with 5 mL Buffer ELU. The DNA was precipitated with 3.5 mL isopropanol and vortexed. The solution was then centrifuged at 5,000 x g for 30 min in a Hermle Microcentrifuge at 4°C. The pellet was washed in 2 mL 70% (v/v) ethanol and again centrifuged at 5,000 x g in a Hermle Microcentrifuge for 5 min before removing the supernatant and allowing the pellet to air-dry under sterile conditions. The pellet was then dissolved in 300 μL of

molecular grade H_2O and plasmid concentration was determined using a Nanodrop™Spectrophotometer.

2.3.10 Restriction by *SfiI* of the SOE fragment and pComb3XSS vector

For ligation of the SOE product into the pComb vector, *SfiI* (New England BioLabs) restriction sites were incorporated into the scFv segment via the primers used earlier. The *SfiI* restriction enzyme allows for the unidirectional cloning of the scFv fragment into the pComb3XSS vector by recognising 8 bases that flank 5 random nucleotides (5' GGCCNNNNNGGCC 3'), leaving DNA overhangs.

The reaction was set up as follows:

Component	**Volume**
SfiI	1.8μL (36U)
SOE DNA	1μg
10X NEBuffer	5μL
H_2O	To 50μL

Component	**Volume**
SfiI	1μL (20U)
pComb3XSS vector	2μg
10X NEBuffer	5μL
H_2O	To 50μL

Five x 50μL reactions were set up for each and incubated at 50^{o}C for 5 hours. Samples were electrophoresed on an agarose gel (see Section 2.3.14) to allow separation and purification of restricted DNA from non-restricted DNA and the reaction components. For the pComb vector, a band of ~3kb (vector with the 1.6 kb stuffer removed) was excised from the gel. Stocks of the pComb3XSS vector (containing a stuffer fragment) were generated through the culture of vector-containing *E. coli* cells and subsequent purification of their plasmid (see Section 2.3.18).

2.3.11 Antarctic phosphatase treatment of pComb3X vector

Prior to ligation of each SOE product into the pComb3X vector, the enzyme Antarctic phosphatase (New England Biolabs) was used to remove the 5' phosphate groups from

the *Sfi*I treated vector. The purpose of this was to prevent self-ligation of the vector when treated with DNA ligase. A lower amount of self-ligation reduces the amount of background clones (lacking an SOE fragment) after transformation. The reaction was set up as follows:

Component	**Volume**
Antarctic phosphatase	1μL (5U)
pComb3X vector	26μL (5μg of DNA)
10X Antarctic Phosphatase Reaction Buffer	3μL

The reaction was incubated for 15 min at 37^{o}C and heat inactivated for 5min at 70^{o}C. The reagents used here have no negative effect on the subsequent ligation reactions; hence, ethanol purification was not necessary. Following antarctic phosphatase treatment, the vector was immediately used in ligation reactions.

2.3.12 Ethanol precipitation of DNA

DNA reactions that required concentration (if small volumes were required in a reaction) and/or purification underwent ethanol precipitation by adding a 0.1X volume of 3M sodium acetate, pH 5.2, a 10X volume of ice-cold 100% ethanol and 1 μL glycogen (Ambion). The solution was mixed and stored at -20^{o}C for 2 hours (or overnight). Following incubation, the DNA was pelleted by centrifugation at 20,000 g (Hermle Microcentrifuge) for 15-30 minutes at 4^{o}C in a microfuge tube. The supernatant was carefully removed without disrupting the pellet. The pellet was washed to remove excess contaminants with 1 mL of 70% (v/v) ethanol and centrifuged at 20,000 g for 2 minutes at 4°C. The supernatant was removed carefully and the tube re-centrifuged

briefly for 2 minutes to remove any remaining liquid. The pellet was air-dried at room temperature under sterile conditions and then dissolved in a predetermined (based on the intended use) volume of molecular grade water.

2.3.13 Quantitation of nucleic Acids and proteins using a Nanodrop 1000

Rapid quantitation of nucleic acid or protein concentration in a solution was determined by measuring the absorbance of at a specific wavelength using a Nanodrop 1000 (Thermo Scientific). The platform is initialised by adding molecular grade water to the pedestal for calibration before blanking using the solution in which the nucleic acid or protein is suspended in. The sample to be quantified (2 μL) is pipetted onto the pedestal of the machine, which is cleaned following each read. Samples are analysed in triplicate to determine sample concentrations. The platform operates on the basis that DNA absorbs light at 260nm and can, therefore, be quantified based on the amount of light absorbed. Protein quantitation is based on light absorbance at 280nm by aromatic amino acids, primarily Tryptophan.

2.3.14 DNA separation by agarose gel electrophoresis

DNA samples can be separated based on size by agarose gel electrophoresis. Gels are prepared by dissolving the desired concentration of agarose in 1X Tris Acetate EDTA (TAE) buffer, pH 8.3 (from a 10X stock solution (Sigma)). Gel concentrations were typically 1-2% (w/v) depending on the expected size of the DNA (concentration increases with DNA band size). Agarose is dissolved in TAE buffer, by heating in a microwave, until the solution is clear. The fluorescent SYBR® Safe DNA gel stain is then added to the solution (1μL per 10 mL). The solution is then poured into gel casts and a comb is placed

penetrating the surface to serve as wells once the gel solidifies. For loading of large volumes of DNA samples, the teeth of the comb can be joined together using masking tape or autoclave tape. The gel is then submerged in 1X TAE buffer in an electrophoresis bath (BioRad Wide Mini Sub Gel GT).

DNA HyperLadder™ (Bioline) (~12μL) is loaded into the first well of the gel to indicate the size of any DNA bands. DNA samples are loaded into the adjacent wells (volume depending on expected DNA concentration). A current is then applied to the gel (typically 90V for 45 min). As DNA is negatively charged it migrates from the wells towards the positive end of the electrophoresis bath (the distance traveled depending on DNA size, agarose concentration, voltage applied and time). DNA is visualised on a Geldoc EZ System Image Lab or a Darkreader.

2.3.15 Ligation of the SOE fragment into vector DNA

The gel-purified, *Sfi*I-restricted SOE fragments from chickens ABG055 and AB048 were cloned into the gel-purified, *Sfi*I-restricted, antarctic phosphatase treated pComb3X vector using T4 DNA ligase (New England Biolabs). Multiple small-scale ligation reactions were carried out to generate large quantities of DNA. Each reaction was a tenth of a full-scale ligation and transformation in size and included a negative control excluding SOE DNA to evaluate the background ligation. The small-scale reaction was set up as follows:

Component	**Quantity**
SOE DNA	70ng

Vector DNA	140ng
10X T4 DNA Ligase Buffer	2μL
T4 DNA Ligase	1μL
H_2O	To 20μL

Note: Concentrations of DNA differed for ABG048 and ABG055 DNA; therefore, volumes used were not consistent. Instead DNA quantities were measured in nanograms (ng). In all cases the reaction volume was 20μL.

The reaction was incubated overnight at room temperature before purification by ethanol precipitation (see Section 2.3.12) and resuspension in 10μL H_2O.

2.3.16 Library transformation into electrocompetent *E. coli* XL1-Blue cells and precipitation of scFv-displaying phage

For the transformation procedure, the ligated DNA samples were stored on ice along with commercial electrocompetent *E. coli* XL1-Blue cells (Agilent Technologies). Electroporation cuvettes (0.2 cm, Bio-Rad) were placed on ice one hour before transformation. Super Optimal broth with Catabolite repression (SOC), a nutrient-rich medium designed for growth following transformation, was warmed to 37^{o}C before the transformation step.

Previously ligated DNA (see Section 2.3.15) was split into two aliquots. Each aliquot was added to a commercial vial of *E. coli* XL-1 Blue electrocompetent cells (300μL) to ensure there was a surplus of cells present during the transformation. The DNA/cell preparation was gently mixed and left on ice for 1 min. An additional preparation of the pComb

vector, enzymatically treated identically to the experimental DNA but in the absence of SOE DNA, was included in the transformation experiment. This was used as a control for self-ligated vector. The DNA/cell mixtures underwent electroporation (Bio-Rad Gene Pulser Xcell^{TM}) with settings of 25 μF, 2.5 kV and the gene pulse controller set at 200 Ω. The cuvettes were immediately flushed with 1 mL SOC medium, followed by 2mL SOC (x2) and transferred to 50mL polypropylene tubes containing 5 mL prewarmed SOC medium. The tubes were incubated for 1 hour at 37^{o}C with shaking at 250rpm. Following this incubation, serial dilutions were set up (10^{-1} to 10^{-8}) and 100μL of each dilution was plated onto LB agar plates (with 100μg/mL carbenicillin) and incubated overnight at 37^{o}C to determine library size. Non-transformed *E. coli* XL-1 Blue cells (10 μL) were plated, onto a separate plate as a control. Following this, 5mL of SB media was added to the culture with 30μL of 5mg/mL tetracycline and 3μL of 100mg/mL carbenicillin. This was left to incubate for an additional hour under the same conditions before addition of 4.5μL of 100mg/mL carbenicillin and then left to incubate for a further hour. This culture, along with 92.5 μL 100 mg/mL carbenicillin, 370 μL of 5mg/mL tetracycline and 2mL of commercial M13KO7 Helper Phage (Invitrogen), was then transferred to a 1L sterile polypropelene flask containing 183 mL SB warmed to 37^{o}C. The culture was incubated stationary for 30 min at 37^{o}C, and then for 90 min at 37^{o}C while shaking at 300 rpm. Kanamycin (50mg/mL) (280μL) was added to the culture to induce phage expression.

The following morning, the culture was divided between 4 x 85mL Oakridge tubes and centrifuged at 3,220 g for 15 min at 4^{o}C in an Eppendorf 5810 benchtop centrifuge. The phage-containing supernatant was removed and 8 g of PEG-8000 and 6 g of NaCl were added for phage precipitation. The solids were dissolved through shaking and the

solution was divided into 4 x 85mL Oakridge tubes to incubate on ice for 30 min. The tubes were then centrifuged at 15,000 g in an Eppendorf 5810 benchtop centrifuge for 15min at 4^{o}C and the supernatant discarded. The remaining liquid was removed through inverting the tube and using a paper towel to collect any 'run-off', leaving only phage pellets. The pellets were resuspended in a total of 4 mL PBS 1% (w/v) BSA and centrifuged at 15,000g in an Eppendorf 5810 benchtop centrifuge for 5 min at 4^{o}C. The supernatant was passed through a 0.2μM filter into 2mL microcentrifuge tubes and sodium azide was added to a final concentration of 0.02% (v/v) for storage at 4^{o}C prior to panning. This solution was considered an unpanned ABG055 scFv-phage library. See section 2.3.19 for panning protocol.

2.3.17 Sequencing of scFv genes and antigen genes

Two different sequence service providers were used during the duration of this work. Sequences generated in Chapter three utilised Source Biosciences as the supplier. Sequences generated in Chapter five utilised GATC Biotech (see Section 2.5.1.2). Sequencing from Source Bioscience required 'in-house' isolation of the plasmid prior to dispatch (see Section 2.3.18).

ScFv genes were sequenced using *ompseq* and *gback* primer combinations. The *ompseq* primer targets a region of the ompA leader sequence while the *gback* primer targets the gene III fragment – flanking both SfiI sites of the scFv fragment. The heavy and light chain complementarity determining regions (CDRs) of the antibodies were identified using the Kabat rules for antibody CDR regions and compared for library diversity.

Ompseq (sense)

5' - AAG ACA GCT ATC GCG ATT GCA G - 3'

Gback (reverse)

5’ - GCC CCC TTA TTA GCG TTT GCC ATC - 3’

2.3.18 Isolation of plasmids from bacterial cells

Plasmids were isolated from single colonies using the NucleoSpin® Mini Prep Kit (Macherey-Nagel). Briefly, single colonies were picked from LB agar plates. These single colonies were incubated overnight in 10mL SB containing the appropriate antibiotic (100 μg/mL Carbenicillin was used for pComb vector-containing cells) at 37°C, 220rpm on an orbital shaker. On the following morning, cultures were pelleted by centrifugation at 3,220 g (Eppendorf Centrifuge 5810 R). The supernatant was discarded and pellets were resuspended in 250μL of kit buffer A1. Kit buffer A2 (250μL) was added and contents were mixed by gentle inversion and left for 5 min at room temperature. Kit buffer A3 (300μL) was added and contents were mixed as before. Solutions were centrifuged at 11,000g for 5 min in a Hermle Microcentrifuge and the supernatant was passed through a NucleoSpin® Plasmid Column into a Collection Tube (2 mL) (11,000 g). The flow-though was discarded and the column membrane was washed by the addition of 500μL kit buffer AW and centrifugation at 11,000g in a Hermle microcentrifuge for 2 minutes. The DNA was precipitated with the addition of 600μL kit buffer A4 and the column was centrifuged at 11,000 g in a Hermle microcentrifuge for 2 minutes. The membrane was then dried for 2min by centrifuging at 11,000 g in a Hermle microcentrifuge before the captured plasmids were eluted by the addition of 50μL Buffer AE and captured through a final centrifugation step at 11,000 g in a Hermle microcentrifuge for 2 minutes.

2.3.19 Enrichment of ABG055 scFv-phage library through cell-based panning

The ABG055 scFv-phage library generated was panned against live PC3 cells using the following protocol.

An overnight culture (50μL) of *E. coli* XL1 blue cells was subcultured into 50mL SB with 100μL of 5mg/mL tetracycline and incubated at 37°C, while shaking at 220rpm, until an OD_{600} of 0.6 was reached. ABG055 scFv-phage unpanned library (1.5mL) was added to the culture and left to incubate static at 37°C for 30min. The culture was pelleted in 2 x 50mL polypropylene tubes by centrifugation at 3,220 g for 15min in an Eppendorf 5810 benchtop centrifuge. The pellets were resuspended in a total of 1.5mL of SB and spread on LB agar 100μg/mL carbenicillin plates (100μL per plate) and left to incubate overnight at 37°C. The following day the plates were scraped into SB and pelleted by centrifugation at 3,220 g for 15min in an Eppendorf 5810 benchtop centrifuge. Pellets were resuspended in 5mL of SB 40% (v/v) glycerol and made into 5 x 1mL aliquots. Two of these were frozen at -80°C for long-term storage. The remaining mixture was split into two and added to two 500mL polypropelene flasks, each containing 100mL SB, 50μL of 100mg/mL carbenicillin and 196μL of 5mg/mL tetracycline. Cultures were propagated at 37°C, while shaking at 220rpm until an OD_{600} of 0.6 was reached, followed by the addition of 1mL of helper phage per 100mL of culture. Cultures were left static for 30min at 37°C and then shaken at 300rpm for 1 hour at 37°C. Kanamycin (140μL of 50mg/mL) was added per 100mL culture and cultures were left overnight at 30°C, while shaking at 300rpm.

The following morning the phage culture was centrifuged at 15,557 x g in an Eppendorf 5810 benchtop centrifuge for 15 min at 4°C. The phage-containing supernatant was

removed and PEG-8000 (4%,w/v) and NaCl (3%,w/v) were added for phage precipitation. The solids were dissolved through shaking and the solution was dispensed into Oakridge tubes and incubated on ice for 30min. The tubes were then centrifuged in an Eppendorf 5810 benchtop centrifuge at 15,557 x g for 15min at 4^{o}C and the supernatant discarded. Any remaining liquid was removed through inversion of the tube and using a paper towel to collect any 'run-off', leaving only phage pellets. The pellets were resuspended in a total of 2mL of PBS with 1% (w/v) BSA (per 100mL of culture) and centrifuged at 15557 x g for 5min at 4^{o}C in a Hermle Microcentrifuge. Supernatant was passed through a 0.2μM filter into a 2mL microcentrifuge tube. Sodium azide was added to a final concentration of 0.02% (w/v) for storage at 4^{o}C. This solution was used as Input phage. Input phage (1mL) was added to 3mL PBS containing 5% (w/v) Milk Marvel, mixed and left on ice for 1 hour.

A starter culture was set up the day before. This culture consisted of *E. coli* XL-1 Blue cells in SB (10mL) with carbenicillin (100μg/mL) and was incubated at 37^{o}C, while shaking at 220rpm to provide cells for subculture. On the following morning, it was subcultured at different dilutions ranging from 1 in 100 to 1 in 5,000 into fresh SB Carbenicillin 100μg/mL (10 mL) and allowed to grow at 37^{o}C, while shaking at 220rpm, to be used later in the day. When phage solutions were ready for cell infection, the absorbance of all cultures (OD_{600}) was measured and the culture with an OD_{600} of at least 0.6 (mid-log phase) was selected for infection. The remaining cultures were not used. Adherent PC3 cells were detached from T175 flasks using 'enzyme-free' dissociation buffer and cells were enumerated using trypan blue (see Sections 2.2.10 and 2.2.6). The number of cells used (see Table 2.2) were centrifuged at 750 g in a Sigma benchtop centrifuge for 5 min and resuspended in 1mL of nutrient supplemented RMPI 1640 (see

Section 2.2.3) at the desired concentration. One mL of suspended cells was added to the 4 mL solution of input phage/milk marvel and gently mixed on a roller-shaker at room temperature for 30 min. The cell/scFv-phage mixture was centrifuged at 750 x g (brakes set on low) for 3 minutes in an Eppendorf 5810 benchtop centrifuge and the supernatant removed. The pellet was resuspended in 180μL FACs buffer in a new centrifuge tube for washing by centrifugation. The number of wash steps performed depended on the round of panning (see Table 2.2).

Following wash steps, the cell-phage pellets were resuspended in 1mL of PBS in a new microcentrifuge tube and centrifuged for 3 min at 750 g in a Hermle microcentrifuge. Supernatant was removed and the pellet was resuspended in a 150 μL solution of 10 mg/mL trypsin in PBS warmed to 37^{o}C. This mixture was incubated for 30 min at 37^{o}C while shaking at 300rpm to elute any bound phage. The solution was centrifuged at 20,000g in a Hermle microcentrifuge to pellet cell debris and the supernatant removed as output phage. The output phage aliquots were combined and added to 4 mL of *E. coli* XL1 Blue cells in mid-log phase and left static for 15 min at room temperature. The cells were pelleted by centrifugation at 3,220 g in an Eppendorf 5810 benchtop centrifuge and resuspended in 500 μL SB. This output library was then spread onto 5 x LB agar 100μg/mL carbenicillin plates (100μL per plate) and left overnight to incubate at 37^{o}C.

The following morning, 3 mL of SB was added to each plate and the 'carpet-like' colonies were resuspended through scraping and pelleted by centrifugation at 3,220 g for 20 minutes in an Eppendorf 5810 benchtop centrifuge. The pellet was resuspended in 5 mL SB, 3 mL of which was added to 96mL prewamed SB with 50 μL 100 mg/mL carbenicillin and 196μL of 5 mg/mL tetracycline, in a 500mL polypropylene flask, for incubation at

37°C, while shaking at 220 rpm, until an OD_{600} of 0.6 was reached. The remaining 2 mL of infected *E. coli* XL1 Blue cells had glycerol added to a concentration of 40% (v/v) and these were then stored as stocks at -80°C. Once the culture reached mid-log phase, 1mL of helper phage was added and left static at 37°C for 30 min, and then shaken at 300rpm for another hour. Kanamycin (140μL of 50mg/mL) was then added and the culture was left overnight at 30°C, while shaking at 300 rpm. The following morning phage were precipitated as before and the next round of panning was initiated using the same protocol. Efficacy of the procedure was determined by scFv-phage polyclonal ELISA (see Section 2.3.21).

Table 2.2 - Panning protocol used to enrich the ABG055 scFv-phage library. Each round increased washing stringency.

Round	Washes	Cell Number
1	4	~2 x 10^7
2	6	~2 x 10^7
3	8	~2 x 10^7
4	10	~2 x 10^7

2.3.20 Colony pick PCR

During certain rounds of panning, colony pick PCRs were carried out to ensure that the scFv fragment was still present in the vector. Single colonies were picked from LB agar carbenicillin (100ug/mL) plates upon which phage-infected *E. coli* XL1 Blue cells had been grown overnight. A sterile pipette tip was used to pick single colonies from a plate and resuspended into 10 μL of 'molecular-grade' water. The PCR-amplified fragments were loaded onto a 1% (w/v) agarose gel, run (90V, 45 min) and visualised. Gels were

investigated for the presence of scFv genes. Amplification was performed under the following conditions:

Component	**Concentration**
CSC-F primer	50μM
CSC-B Primer	50μM
Suspended single colony	10μL
MyTaq Red Polymerase (5U/μL)	1μL
5x MyTaq Red Buffer	10μL
Molecular Grade Water	To 50μL

The PCR program utilised was the same as for SOE generation (see Section 2.3.8).

2.3.21 Polyclonal ELISA of panned phage-scFv

To test output scFv-phage for enrichment against PCa cells, a polyclonal phage ELISA was carried out. PC-3 cells were seeded into a Nunc 96-well cell culture plate (10,000 cells per well) in 200 μL media (see Sections 2.2.4 and 2.2.6) and left to incubate at 37^{o}C, 5% CO_2 for ~18 hours to allow for cell adherence and growth. The following day the medium was removed and 100 μL of 8% (v/v) formalin was added to each well (for fixation) for 25 min at 37^{o}C. The plate was washed three times with dH_2O, then three times with PBS. Each well was blocked with 150 μL 5% (w/v) Milk Marvel in PBS for 1 hour at 37^{o}C. Phage eluted from each round were diluted threefold in 5% (w/v) Milk Marvel in PBS in 55μL per well aliquots and left to incubate at room temperature for 1 hour. The blocking solution was removed from the plates and wells were washed three

times with PBS. The phage solutions were added to each well and incubated at 37°C for 1 hour. The plate was washed x3 with PBST then x3 with PBS. HRP-conjugated αM13 diluted 1 in 5,000 in 1% (w/v) milk marvel in PBS (50μL) was added to each well and left to incubate for 1 hour at 37°C. The plate was washed x3 with PBS and PBS-T. TMB (50μL) was added per well and left to develop in the dark at room temperature for ~30min before being neutralised by addition of 50μL of 10% (v/v) HCl. The absorbances were read at 450nm on a Tecan Safire2TM plate reader.

2.3.22 Preparation of soluble antibody fragments

For soluble expression of scFv that underwent enrichment through panning, Top10F' cells were grown in LB tetracycline (5 μg/mL) overnight (37°C, while shaking at 220 rpm) and subcultured into SB tetracyline (5 μg/mL) the following day. At an OD_{600} of 0.6, 2 separate 3 mL aliquots were infected with 20 μL output phage from panning rounds 3 and 4, respectively. Following a 30-minute incubation at 37°C, while shaking at 220rpm, the cells were serially diluted and plated in 100 μL aliquots on LB agar Carbenicillin (100μg/mL) plates and incubated overnight at 37°C.

The following morning, 4 x 96 well plates were prepared (under sterile conditions) with 200 μL SB, 100 μg/mL Carbenicillin and 1% (w/v) glucose per well. Single colonies were picked and added to each well (288 from round 3; 96 from round 4) and grown overnight at 37°C, while shaking at 220rpm. They were then subcultured into deep-welled plates containing 1mL SB, 100μg/mL carbenicillin and 1X 505 (0.5% (v/v) glycerol, 0.05% (v/v) glucose final concentration) and grown at 37°C, while shaking at 220 rpm, to reach an OD_{600} of ~0.6. Following this, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added

to a final concentration of 1mM to each well and plates were incubated overnight at 30^{o}C, while shaking at 220 rpm, in an orbital shaker. The original picked clones (used for subculture) were made into master stock plates by adding glycerol to a concentration of 40% (v/v) and storing them at -80^{o}C.

The following morning the Top10F' culture plates were pelleted at 3,330 x g in an Eppendorf 5810 benchtop centrifuge for 25 min and resuspended in 1 mL PBS per well. The cultures then underwent 3 freeze-thaw cycles of freezing at -80^{o}C (30 minutes) and thawing at 37^{o}C (45-60 minutes) to release scFv from the periplasm. The lysate was diluted 1 in 2 to give a final volume of 2 mL scFv lysate 1% (w/v) BSA in PBS solution, also containing dissolved 'Complete™ Ultra protease inhibitor tablets (Roche) diluted to a quarter of their manufacturer-recommended working concentration. The plates were stored at -80^{o}C until used for screening experiments.

2.3.23 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Electrophoresis is applied to separate proteins based on their size. Originally described in 1970 (Laemmli, 1970), the method utilizes the anionic detergent sodium dodecyl sulphate (SDS) to denature and apply a net negative charge (over a wide pH range) to polypeptide chains. This approach introduces an electric field to cause the negatively charged polypeptide molecules to migrate towards an anode (positively charged). Introducing a matrix in the form of a polyacrylamide gel restricts this migration. The SDS, following denaturation, binds and applies a negative charge to proteins based on their mass. The matrix-restrictive force varies based on the protein's size, meaning smaller

proteins will migrate faster than a larger protein through the gel. The charge applied to each protein by the SDS is almost entirely dependent on the size (Schägger and von Jagow, 1987). Hence, the technique is excellent for size-based separation of proteins in complex solutions. Introducing a protein ladder of known size also allows the user to estimate the size of proteins treated with SDS-PAGE. For SDS-PAGE, a gel consisting of two stacked polyacrylamide matrix concentrations (12.5% and 4.5%) was made up between two glass casting plates. Gel constituents can be found in section 2.1.7.3. First the lower portion of the gel was made, leaving the addition of the APS and TEMED (oxidizing agents responsible for acrylamide polymerization) as the final step, and poured between the glass casting plates. A thin layer of ethanol was added between the casting plates to ensure an even surface of the 12.5% layer. Once the lower layer had set, the ethanol was removed through inversion of the casting plates. As with the 12.5% gel, 4.5% stacking gel constituents were mixed and added on top of the already set 12.5% gel. A plastic comb was embedded into the 4.5% stacking gel prior to oxidization to create protein loading wells in the final gel.

Upon completion of polymerisation of the 4.5% gel, the gel and its accompanying glass plates were placed in to an electrophoresis chamber. The chamber was filled with electrophoresis buffer (25 mM Tris Base, 250 mM glycine, 0.1% (w/v) SDS, pH 8.3) and the gel's comb removed – filling the wells with buffer. Protein samples to be analysed were prepared by adding 4x loading dye and (if the protein concentration was considered too high) molecular grade H_2O to a concentration of 1x. The total volume of protein/dye solution added to each was typically 20 μL. Page Ruler Plus Protein Ladder (Fisher) (4 μL) was added to flanking wells to allow for size determination of proteins when visualised. Samples were typically run at 110V until the bromophenol blue tracker

dye (present in the protein loading dye) reached the bottom of the gel (~90 min). The gel was then removed and rinsed with dH_2O before being incubated with InstantBlue™ on a platform shaker under gentle agitation for 10-30 min. Following protein staining, residual InstantBlue™ dye was removed and the gel was rinsed with dH_2O. Images of gels were taken using a 20MP Zeiss PureView Lumia Camera.

2.3.24 Western blotting

Western blotting is a common immunoassay technique used to determine the presence or absence of a molecule of interest in a complex protein solution such as serum or a cell lysate. Samples must first be separated via SDS-PAGE (see Section 2.3.23) prior to investigation by western blotting. Following separation of protein samples by SDS-PAGE, the resultant gels are incubated in western blotting transfer buffer (Thermo Scientific™ Pierce™ 1-Step Transfer Buffer) for 10 min. In addition to the gel, four sheets of absorbent Western blotting paper and one sheet of a nitrocellulose membrane (all cut to the same dimensions as the gel) are soaked in transfer buffer also for ten minutes. Following buffer incubation, the constituents are set up in a 'sandwich-like' fashion from bottom to top: absorbent paper x 2, nitrocellulose membrane, SDS-PAGE gel and absorbent paper x2. A manufacturer-provided blot roller was used to remove any bubbles between the gel and the nitrocellulose membrane prior to addition of the final two layers of absorbent paper. For transfer of proteins from the SDS-PAGE gel to the nitrocellulose membrane, the 'sandwich' is set-up in the centre of an anode plate cassette of a Thermo Scientific™ Pierce™ G2 Fast Blotter (see Figure 2.2). The cassette was placed into its parent apparatus (Thermo Scientific™ Pierce™ G2 Fast Blotter) and the separated proteins were transferred to the nitrocellulose membrane typically at Mixed Range MW (25-150kDa) setting (11V, 1.3A for 7 min).

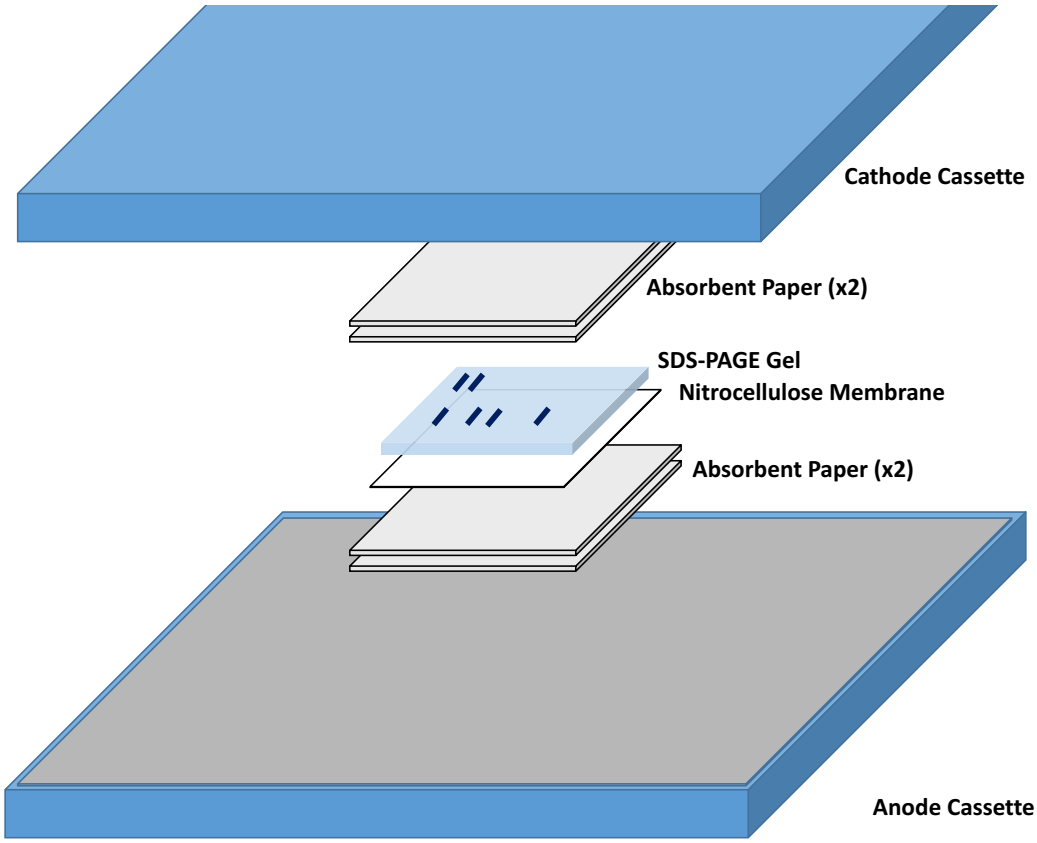

Figure 2.2 - Diagram of sandwich set-up required for Western Blot using the Thermo Scientific™ Pierce™ G2 Fast Blotter system.

Following transfer, the nitrocellulose membrane was blocked in 5% (w/v) Milk Marvel in PBS typically overnight at 4^{o}C, but also at room temperature for two hours. Following completion of blocking, the membrane was washed three times in PBST for 5 min, followed by washing three times in PBS for 5 min. The membrane was then incubated with primary antibody (dilution factor varied based on antibody used) in 1% BSA (w/v) PBST under agitation at room temperature. The membrane was washed as before and incubated with secondary antibody (dilution factor varied based on antibody used) in 1% BSA (w/v) PBST for 1 hour at room temperature under agitation. The combination of secondary antibodies used in the western blots varied based on the assay, though the final antibody used to probe the array was always conjugated to a horseradish peroxidase enzyme for colorimetric development of the substrate 3, 3', 5, 5' -

tetramethylbenzidine (TMB). Prior to development, the membrane was washed three times in PBST and three times in PBS as before. For development, (TMB) Liquid Substrate System for Membranes (Sigma) was added to the entire membrane using a Pasteur pipette. The colourless substrate was converted to a permanent insoluble blue/purple precipitate where HRP was present on the membrane, indicating the qualitative presence of the protein of interest. The reaction was stopped by rinsing the membrane with dH_2O, once colour development was detected. Images of blots were taken using a 20MP Zeiss PureView Lumia Camera.

2.4 High density protein Arrays for monitoring processes in recombinant antibody library generation

2.4.1 Western blot visualisation of an avian immune response to PCa cells

To visualize the chicken AGB055's immune response to a PCa immunisation regime by western blotting, a preparation of PC3 cell lysate was first obtained (see Section 2.2.9). The protein concentrations of PC3 cell lysates were quantified using a Nanodrop™ 1000, following which multiple lanes containing 50 μg of cell lysate was separated by SDS-PAGE (see Section 2.3.23) and transferred to a nitrocellulose membrane for Western Blotting and blocked overnight at 4^{o}C (see Section 2.3.24). One replicate of the SDS-PAGE gel was stained using InstantBlue™ to ensure separation of the lysate. The following morning the membranes were washed in PBST for 5 min (x3) and PBS for 5 min (x3). IgY-containing serum isolated from the chicken ABG055, before and after PC3 and LNCaP C4-2 cell immunization regimes, were diluted 1 in 1,000 in 1% (w/v) MM in PBST respectively. Separate lysate-containing membranes were incubated in 5 mL of pre- or post-bleed serum-containing solutions for one hour on a roller shaker at room temperature. The membranes were washed as before (three times with PBST and three times with PBS) and each was incubated with 5 mL rabbit anti-chicken IgY (Fisher) (diltuted 1 in 5,000) in 1% (w/v) MM PBST on a roller shaker for 1 hour at room temperature. The membranes were then washed as before and incubated with peroxidase conjugated goat anti-rabbit IgG (Sigma) (1 in 1,000) in 1% (w/v) MM PBST for 1 hour at room temperature on a roller shaker. After incubation, both membranes were washed as before and developed in parallel with TMB. Once substrate conversion was visualised both membranes were rinsed, in parallel, with dH_2O to stop the reaction.

2.4.2.1 UniPEx high density protein arrays for immunoassays

High Density UniPEx imaGene Protein Arrays were previously acquired from Source Bioscience imaGenes (Berlin, Germany) by Dr. Gregor Kijanka. These arrays consist of two 22 x 22 cm Polyvinylidene fluoride (PVDF) membranes with 15,300 robotically-spotted *E. coli* colonies chemically fixed to their surface. These colonies were isolated from several human cDNA libraries, isolated from multiple tissues. The UniPEx library used here was generated from 100,000 manufacturer-sequenced clones; isolated from human foetal brain, T-cell and lung libraries (Büssow *et al.*, 1998). Source tissue cDNA was cloned into variants of the IPTG-inducible pQE vector (pQE30NST and pQE80LSN) (Source Bioscience imaGenes, Germany). Only in-frame Open Reading Frame clones from the previously mentioned libraries were sequenced and spotted onto the membranes yielding a total of 7,390 distinct human proteins (Source Bioscience, 2010). Prior to fixation, clones were grown directly on the membranes and induced to express their encoded His-tagged human proteins. Global protein expression of human proteins on the arrays was normalised by the manufacturer, before fixation.

2.4.2.2 Activation of protein array surfaces

Arrays were supplied on dry hydrophobic PDVF membranes that required activation to allow for aqueous buffer interaction, before introducing protein-containing solutions to their surface. To achieve this, both membranes (library is spotted on two separate membranes) were placed separately in petri-dishes of the same dimensions. Arrays were rinsed and incubated in 20 mL of 70% (v/v) ethanol for 5 min (volume was sufficient to ensure complete wetting of array). The ethanol was removed and both arrays were rinsed twice with 20 mL molecular grade water. Arrays were transferred into separate

boxes containing 25 mL Tris-Buffered Saline TritonX-100 (TBST-T) (20 mM Tris Base, pH 7.5, 0.5 M NaCl, 0.1% (v/v) Tween 20, 0.5% (v/v) TritonX-100). *E. coli* colonies were wiped from the surfaces using sterile (Kintech W4 Crew® wipers (Sigma), taking care not to contact the surface with gloves etc. This step was repeated with fresh TBST-T to make sure no colonies remained on the surface. Both arrays were washed in 40 mL TBST (20 mM Tris-Base, pH 7.5, 0.5 M NaCl, 0.05% (v/v) Tween 20) on a platform shaker at 12 rpm for ten minutes (x3) and then in 40 mL TBS (20 mM Tris-Base, pH 7.5, 0.5 M NaCl) for 5 min (x2). The arrays were then blocked in 40 mL 5% (w/v) MM in TBST for 2 hours at room temperature or overnight at 4^{o}C (depending on the proceeding immunoassay) on a platform shaker (Stuart Scientific) at ~12 rpm. Following completion of blocking, the arrays were ready for use (provided they have not been allowed to dry).

2.4.3 Protein array profiling of secondary antibodies

Before using protein arrays to profile antibody solutions, it was necessary to first characterise the secondary antibodies selected for each experiment. Any antigens bound by secondary antibodies in the absence of their intended primary antibody could subsequently be excluded from antigen lists as 'false-positives'. A schematic of the protein array characterisation of complex antibody solutions can be seen in Figure 2.3. In the case of IgY profiling, the protocol detailed in section 2.4.4 was applied. The addition of chicken sera to the protein arrays, however, was skipped. instead the protocol used the secondary antibodies rabbit anti-IgY antibody and an alkaline phosphatase-conjugated goat anti-rabbit IgG antibody in the absence of sera. For profiling the secondary antibodies selected for detection of array-bound scFv-phage (rabbit anti-M13 and fd phage coat proteins antibody coupled with an alkaline phosphatase-conjugated goat anti-rabbit IgG antibody), the protocol is detailed in

section 2.4.5 was carried out, with the exception being that all sets introducing scFv-phage to the immunoassay were excluded. Finally, the antibodies selected to detect protein array-bound scFv (a rabbit anti-HA antibody coupled with an alkaline phosphatase conjugated goat anti-rabbit IgG antibody) were characterised on the array using the protocol described in section 2.4.6, without introducing scFv into the immunoassay.

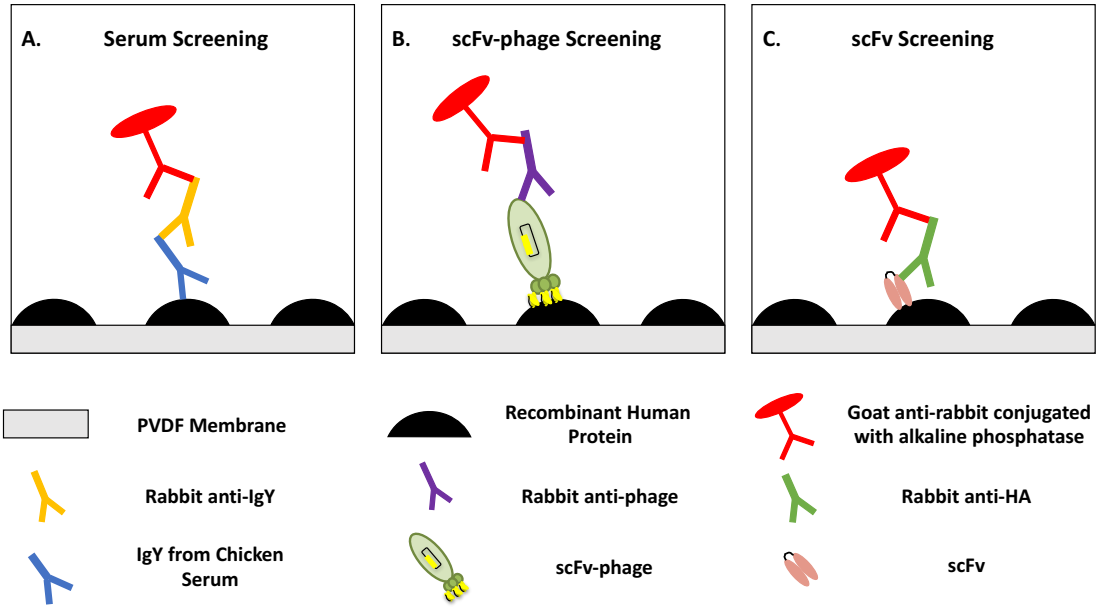

Figure 2.3 - Summary of Immunoassay formats used in UniPEx Protein Array Profiling experiments. Protein Arrays were used to profile: A. serum isolated from a PCa cell line-immunised chicken before and after immunisation, B. an scFv-phage library generated from a PCa cell line-immunised chicken, and C. scFv enriched against a PCa cell line via panning technology (each immunoassay is depicted in one of the three above tiles). Prior to profiling experiments, the secondary antibodies selected for each experiment were profiled on a protein array in the absence of their intended antigen. For serum profiling experiments, rabbit anti-IgY antibody (Fisher) and an alkaline phosphatase-conjugated goat anti-rabbit IgG antibody (Sigma) were profiled in the absence of IgY. For scFv-phage screening experiments, rabbit anti-M13 and fd phage coat proteins antibody (Abcam) coupled with an alkaline phosphatase conjugated goat anti-rabbit IgG (Sigma) antibody were profiled in the absence of scFv-phage. For scFv screening experiments, a rabbit anti-HA antibody (Sigma) coupled with an alkaline phosphatase-conjugated goat anti-rabbit IgG antibody was profiled in the absence of scFv. The same goat-anti rabbit with alkaline phosphatase (Sigma) was used in all experiments.

2.4.4 Epitope profiling of an avian immune response to prostate cancer cells using high density protein arrays

For serum profiling experiments, UniPEx Protein Arrays were activated as detailed in section 2.4.2.2. Following a blocking step in 5% (w/v) MM in TBST for two hours at room temperature on a platform shaker, the arrays were washed in 40 mL TBST on a platform shaker at 12 rpm for ten min (x3) and then in 40 mL TBS for 5 min (x2). Two sets of arrays were placed back-to-back in the same container to display the full UniPEx library. Serum (pre- and post-immunisation) obtained from the chicken ABG055 (see Section 2.3.2 for serum isolation details) was diluted 1 in 300 in 40 mL 2% (w/v) BSA in TBST and added to the arrays. Note: optimisation of serum concentration on arrays was carried out in earlier experiments, where it was observed that binding signals were obscured by a higher background in more concentrated serum samples. Pre-immunisation bleed serum was added to one set of arrays while post bleed serum was added to the other. Arrays were incubated in serum-containing solution on a platform shaker (12 rpm) for 1 hour at room temperature, then moved to a cold room to incubate overnight at 4^{o}C on a platform shaker (12 rpm). On the following morning, the arrays were washed as before. Rabbit anti-chicken IgY (Fisher, Product code 31104, Lot code PK19380211) diluted 1 in 1,000 in 2% (w/v) BSA in TBST (40 mL), was added to both sets of arrays and incubated at room temperature on a platform (~12 rpm) shaker for 2 hours. Arrays were washed as before. Following washing, the arrays were incubated for two hours under the same conditions with alkaline phosphatase-conjugated goat anti-rabbit IgG (Sigma, Product code A3687, Lot code SLBJ6146V) at a 1 in 1,000 dilution in 2% (w/v) BSA in TBST. Following incubation, the arrays were washed as before, then incubated for 10 min in AP buffer (1 mM $MgCl_2$, 0.1 M Tris, pH 9.5), followed by 10 min in 25 mM Attophos in AP buffer. The arrays were placed face up on a FujiScanner Fla5100 and scanned for

fluorescence (Attophos has a max excitation of 435 nm and max emission 555 nm. Attophos fluorescence is detactable with the Fla5100 instrument using the preinstalled 473 nm laser for excitation and the Filter LPB (wavelength 513-640 nm) for emission (FujiFilm, 2016)).

2.4.5 Epitope profiling of an anti-PCa scFv-phage library

For scFv-phage epitope mapping experiments, the unpanned scFv-phage library generated from the PCa cell immunised chicken, ABG055, was reamplified and tested against the UniPEx arrays for binding. The library was reamplified on the day of the experiment, as scFv displayed on phage coat proteins can detach when undergoing long term storage (Barbas III *et al.*, 2001).

For reamplification of the scFv-phage library, one 20 mL culture of log-phase *E. coli* XL-1 Blue (in SB with 5μg/mL tetracycline) were incubated statically for 30 min with 1 mL unpanned ABG055 scFv-phage (see Section 2.3.16 for scFv-phage library generation protocol). The infected cells were split and added to two 500 mL polypropylene flasks, each containing 100 mL SB, 50 μL 100 mg/mL carbenicillin and 196 μL of 5μg/mL tetracycline, after the addition of the infected cells. The cultures were grown at 37°C, while shaking at 200 rpm, until an OD_{600} of 0.6 was reached. Helper phage was then added (1 mL per 100 mL culture). The flasks were left static for 30 min then shaken at 300 rpm for 1 hour at 37°C. Kanamycin (140 μL of 50 μg/mL) was added to each flask and the cultures were incubated overnight at 30°C, while shaking at 300 rpm, for phage propagation. The following morning the cultures were centrifuged in Oakridge tubes at 15,557 g in an Eppendorf 5810 benchtop centrifuge for 15 min at 4°C. The phage-containing supernatant was removed and PEG-8000 (4%, w/v) and NaCl (3%, w/v) were

added for phage precipitation. Any solids were dissolved by shaking and the solution was dispensed into Oakridge tubes and left to incubate on ice for 30 min. The tubes were then centrifuged at 15,557 g in an Eppendorf 5810 benchtop centrifuge for 15 min at 4^{o}C and the supernatant discarded. Any remaining liquid was removed through inversion of the tube and using a paper towel to collect any 'run-off', leaving only phage pellets. The pellets were resuspended in a total of 2mL PBS with 1% (w/v) BSA (per 100 mL of culture) and centrifuged at 15557 g in an Eppendorf 5810 benchtop centrifuge for 5min at 4^{o}C and supernatant was passed through a 0.2μM filter into a 2mL microcentrifuge tube. Sodium azide was added to a final concentration of 0.02% (w/v) for storage at 4^{o}C. For addition to the UniPEx arrays, 1 mL of the scFv-phage mixture was added to a 3 mL solution of PBS containing 5% (w/v) MM and incubated on ice for 1 hour.

UniPEx arrays were prepared the day before scFv-phage precipitation according to the protocol in section 2.4.2.2, and left to block overnight at 4^{o}C. The following day, the arrays were washed for 5 min in TBST (x 3) and for 5 min in TBS (x 3). The scFv-phage, in Milk Marvel mixture, were added to 40 mL 5% (w/v) Milk Marvel in TBST and incubated overnight on a platform shaker (~12 rpm) at 4^{o}C. The following morning the arrays were washed as before and rabbit anti-M13 + Fd Bacteriophage coat proteins (2 μg/mL in 40 mL) (Abcam, Product code ab6188, Lot code GR229940-1) in 1% (w/v) BSA in TBST were added to the arrays and left to incubate for two hours at room temperature on a platform shaker (~12 rpm). The arrays were then washed as before and goat anti-rabbit IgG conjugated to alkaline phosphatase (AP) (Sigma, Product code A3687, Lot code SLBJ6146V) at a 1/1,000 dilution in 2% (w/v) BSA in TBST was added to incubate on a platform shaker (~12 rpm) for two hours. Following incubation, the arrays were washed

as before and then incubated for 10 min in AP buffer, followed by 2 min in 25 mM Attophos in AP buffer. The arrays were placed face up on a FujiScanner Fla5100 and scanned for fluorescence (Laser wavelength 473 nm, Filter LPB, Resolution 50μm).

2.4.6 Epitope profiling of a PCa cell-panned scFv library

For epitope mapping of 384 soluble scFv previously enriched against Prostate cancer cells (see Sections 2.3.19 and 2.3.22). Lysates of 384 solubilised scFv were prepared and stored at -80^{o}C. Arrays were activated for immunoassay (see Section 2.4.2.2) and left to block in 5% (w/v) MM (40 mL) for two hours at room temperature prior to the addition of the scFv.

Following blocking, the arrays were washed for 10 min in TBST (x 3), and then for 5 min in TBS (x 2). Deep-well plates containing the lysate of the scFv were thawed in a water bath at 37^{o}C and centrifuged in an Eppendorf 5810 benchtop centrifuge (30 min, 3333 g, 4^{o}C) for ensure the supernatant was clear. Lysate (300 μL) from each well was pooled to give a total volume of 115.2 mL. BSA was added to a concentration of 2% (w/v) and the solution was mixed to ensure uniform dispersion. UniPEx arrays were separated in two petri dishes and the scFv solution was dispensed equally to both for incubation. The arrays were incubated for 1 hour at room temperature, and then overnight at 4^{o}C on a platform shaker (~12 rpm). The following morning samples were washed as before and 40 mL of rabbit anti-HA (Sigma, product code SAB4300603, Lot code 5117T501) diluted 1 in 1,000 in 2% BSA (w/v) TBST was added for a 2-hour incubation on a platform shaker (~12 rpm) at room temperature. The array was washed as before and 40 mL of alkaline phosphatase-conjugated goat anti-rabbit IgG antibody (Sigma, product code A3687, lot code SLBJ6146V) was added for a two-hour incubation under the same conditions. The

arrays were then washed as before and soaked in AP buffer for ten min. The arrays were then transferred into 25 mM Attophos in AP buffer and left to incubate for 2 min before being scanned for fluorescence (Laser excitation wavelength 473 nm, Emission Filter LPB (filter wavelength 513-640 nm), Resolution 50μm) on a FujiScanner Fla5100.

2.4.7 Image analysis

All clones were spotted onto PVDF membranes, in duplicate, in a 3 x 3 pattern. The middle spot in these squares acts as a navigation marker, with clones arranged in one of four patterns around it (see Figure 2.4A). The specific spotting patterns serve as an internal control, so any artefacts present are not mistaken as positives. Only clones detected twice in the same square are defined as positive. Additionally, positive signals were scored manually based on their intensity (with intensity 3 being the strongest and 1 the weakest) (see Figure 2.4B for example of variety of intensities). The quantity of protein spotted on membranes are controlled for consistency during manufacture (Source Bioscience, 2010), so that higher intensities are expected to be due to binding of a high affinity antibody on the array.

A. **B.**

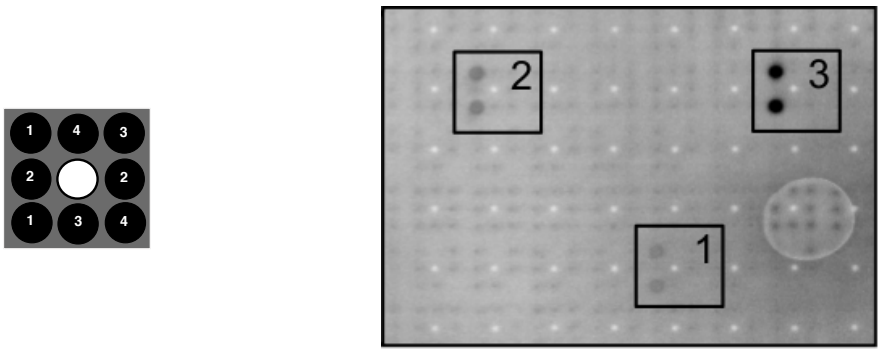

Figure 2.4 – A. The left side of the diagram illustrates the protein spotting placement around navigation/guide dots on the Unipex Array. Here, a single 3 x 3 square of Unipex Array is depicted. Each protein (depicted in black with a white number) is expressed in duplicate around a visible ink guide dot (depicted in white). A total of 4 proteins are spotted in distinctive duplicate patterns around each guide dot. Arrays are scanned for fluorescence using a FujiScanner Fla5100 and images are analysed using VisualGrid software. B. The right side of the diagram is a representative section from a full generated array image, illustrating antibody-antigen binding at three different signal intensities; 3 = strong, 2 = intermediate and 1 = weak, is shown here. In all cases, the proteins identified are spotted in a pattern '1' (depicted in section A of the image, where the duplicate proteins are present in both corners on the left of the grid). The circular artefact on the right-hand side of the image was an air-bubble present between the membrane and the scanner platform at the time of signal retrieval.

Array images were captured using a FujiScanner Fla5100 on settings compatible with the fluorescence of Attophos (Excitation 482 nm, Emission 560 nm – LPB filter). Generated images were analysed using VisualGrid software (GPC Biotech). The software overlays the manufacturer's protein spotting grid on the image generated, matching theoretical protein co-ordinates to their actual location in the image. Due to the size of the array, it was not unusual for background signal to vary in different regions of the

image. The dedicated software also provided the user with the ability to normalize varying background fluorescence in different regions of an image so that positive signals were not obscured during examination (an example of visible positive signals over a normalised background can be seen in Figure 2.5).

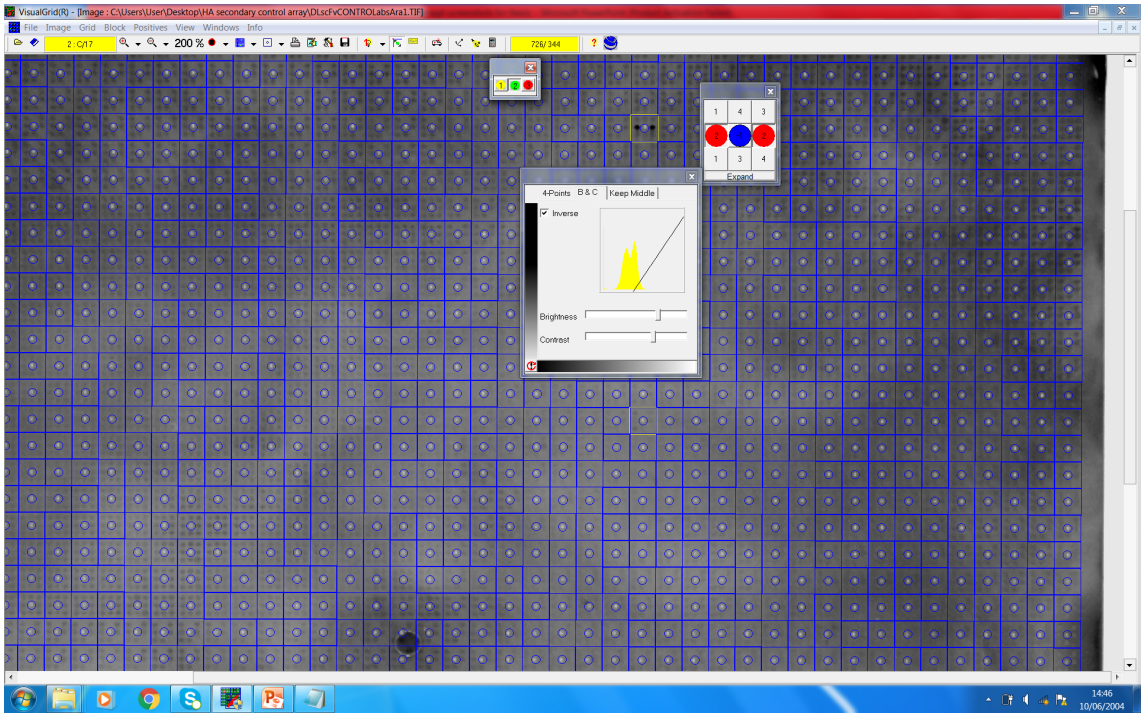

Figure 2.5 - Image analysis was performed using VisualGrid software (GPC Biotech). The software allowed placing a grid (blue lines) on the array image, with yellow boxes marking positives. Software allows for modification of image contrasts and brightness to visualise chemiluminescent signals through background. In this image a positive is identified in a pattern '2' (two horizontal spots flanking a navigation marker) above the image histogram window. In all cases, only clones detected in duplicate were defined as positive.

Upon completion of image analysis using the VisualGrid software, all positives present on the array are marked based on their signal intensity (see Figure 2.6). For the generation of antigen lists, software created by Raymond Moran of the Bioinformatics and Molecular Evolution Group in Dublin City University removed signals from profiling experiments (see Sections 2.4.4, 2.4.5 and 2.4.6) that overlapped with profiles acquired from their corresponding secondary antibodies (see Section 2.4.3) to remove false-

positives. The X.Y positions of the resultant antigen lists were then cross-referenced with a manufacturer-provided database of all array-contained clones - revealing the identity of bound antigens. The database provided contains the manufacturer's clone ID, GenBank ID, Unigene ID and a short description of the identified antigen.

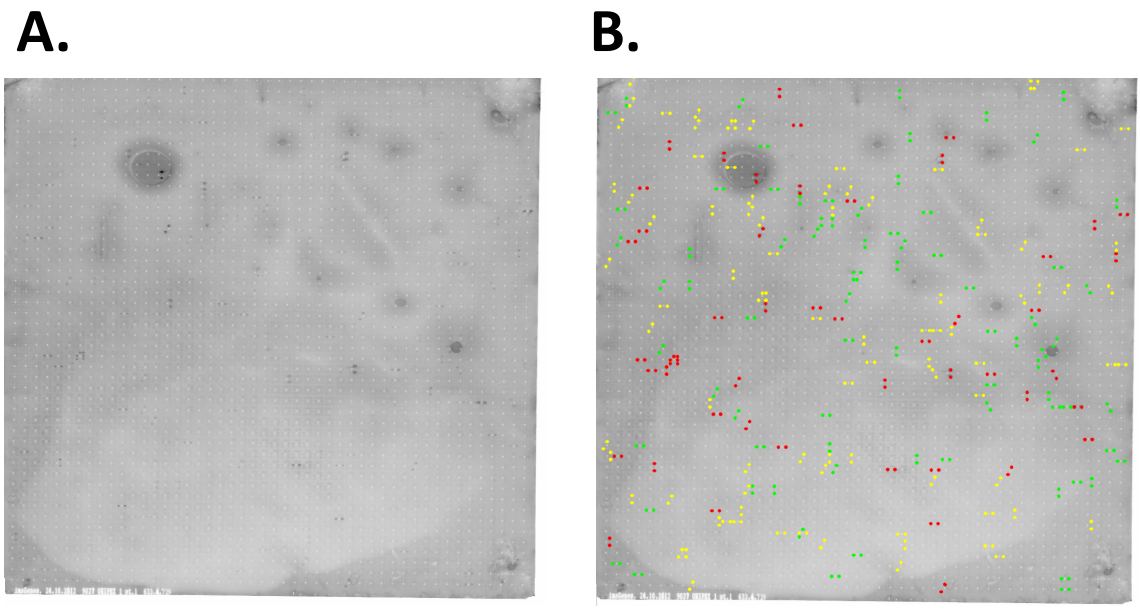

Figure 2.6 – Example of processing protein arrays using VisualGrid software. A - Raw image showing protein array antigen profiling of scFv from 384 clones enriched against PCa cells using cell panning. B – VisualGrid-processed image illustrating all identified binding locations in Image A. Binding signals are depicted as weak (yellow), moderate (green) and strong (red) signals.

2.4.8 BLAST and SIM analysis of protein array-identified antigens of rabbit anti-IgY and alkaline phosphatase conjugated goat anti-rabbit IgG in the absence of chicken sera

To investigate if array protein of rabbit anti-IgY and alkaline phosphatase conjugated goat anti-rabbit IgG secondary antibody binding could be explained by identified antigen similarities with the immunogen used by their manufacturers for antibody production (rabbit IgG and chicken IgY), sequences of identified antigens were obtained via the PubMed website (http://www.ncbi.nlm.nih.gov/protein/) using IDs present in the Unipex protein database. Genes coding segments of the constant and variable regions, of both light and heavy chains, of the species' immunoglobulin were compared with identified antigens. Genes selected for comparisons were chicken immunoglobulin proteins [Ig lambda chain C region (NCBI Accession: P20763.1), Ig lambda chain V-1 region (NCBI Accession: P04210.1), immunoglobulin Y heavy chain constant region (NCBI Accession: XP_015130394.1) and immunoglobulin Y heavy chain variable region (NCBI Accession: ADF29959.1)] and rabbit immunoglobulin proteins [Ig gamma chain C region (UniProtKB: P01870), immunoglobulin heavy chain VDJ region, partial (NCBI Accession: AAA51320.1), Ig lambda chain C region (UniProtKB: P01847.2) and Ig lambda chain variable region, partial (NCBI Accession: AAA31364.1)]. For antigen similarity comparisons, sequences were comparatively analysed for overlap using the Basic Local Alignment Tool (BLAST) for protein-protein overlap (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). BLAST compares the entire length of protein sequences for similarity. However, proteins may also share a few common domains interspersed with regions of little to no homology, that may serve as epitopes to characterised antigens. To investigate for non-intersecting protein sequence

alignments, the local similarity program SIM was used (http://expasy.org/tools/sim-prot.html). SIM divides two protein sequences into smaller sub-sections (in this case proteins were split into 20 AA sequences) and ranks overlapping fragments based on score. Here, the BLOSUM62 matrix was applied for scoring, as it is optimised for comparison of distant proteins (Pearson, 2013).

2.4.9 DAVID and Orthology analysis of array-identified antigen lists

To investigate if immunisation or panning experiments enriched for proteins of specific cellular location, process, function or pathway, the array-acquired antigen lists were compared to the Unipex proteome using the web-based Database for Annotation, Visualization and Integrated Discovery (DAVID) functional annotation tool (https://david-d.ncifcrf.gov/summary.jsp) (Huang *et al.*, 2008). In addition, the same antigen lists were investigated for similarity with their ortholog in chicken, using the Biomart data mining tool (http://www.ensembl.org/index.html). Orthologs are genes in different species that have evolved from the same ancestral gene. These investigations were carried out in collaboration with Raymond Moran of the Bioinformatics and Molecular Evolution Group in Dublin City University.

For DAVID analysis, the newest version of the software was utilised (DAVID 6.8 Beta). DAVID combines public bioinformatic resources into a single database, from which experimental gene lists can be analysed and compared for enrichment to biological functions. The antigen profiles were compared to the complete Unipex proteome to determine if enrichment to antigens, based on cellular location, process, function or pathway had occurred. The Imagene/Genbank IDs of the antigen lists generated from

sections 2.4.4, 2.5.5 and 2.5.6, in addition to the complete Unipex proteome, were converted into their Entrez Gene IDs to utilise the most up-to-date records of each gene (Maglott *et al.*, 2011). The lists were then compared to the Unipex proteome for enrichment using DAVID software. The antigens present in each dataset were clustered by their roles in cellular pathways and functions. Multiple statistical investigations were carried out by the software to provide an overall enrichment score to each pathway and function investigated.

Orthology analysis was carried out to determine if the array-generated datasets were enriched based on homology with their orthologs in chicken. The Imagen/Genbank IDs of each dataset were converted into Ensembl ENST IDs, to allow for compatibility with Biomart. The human and Unipex genomes were also included in the analyses for comparative purposes. The Ensemble-hosted tool was mined for each gene's corresponding ortholog in chicken, as well as the number of orthologs present, the last common ancestor of the gene and the overall percentage homology between the orthologs. The resultant datasets were then compared for qualitative trends.

2.5 Screening for protein array-identified antigens of an avian scFv library

2.5.1.1 Expression and purification of PSMA7, Retinoic Acid Receptor Gamma, SET Domain-Containing 2 and ZNF358

E. coli expression clones encoding for recombinant PSMA7 (IMGSp9027C1115D), Retinoic Acid Receptor Gamma (IMGSp9027C0873D), ZNF358 (IMGSp9027B1237D) and SETD2 (IMGSp9027C0934D) were acquired as stab cultures from Source Bioscience (Cambridge, UK). The proteins were supplied cloned into the pQE30NST vector (Bussow *et al.*, 1998). The pQE-derived vector contains a T5 promotor followed by a double lac operon insert upstream of the encoded protein (inserted into the vectors multiple cloning site) with an RGS-His_6 tag to allow for IPTG-induction of expression and purification via metal affinity chromatography (IMAC).

Upon receipt, the clones were spread on LB agar Ampicillin (100 μg/mL) and Kanamycin (15 μg/mL) plates with 2% (w/v) glucose, and incubated overnight at 37°C. Single colonies were picked from the plates and inoculated separately into 10 mL of 2YT medium with 2% (w/v) glucose, Ampicillin (100 μg/mL) and Kanamycin (15 μg/mL) overnight at 37°C, whilst shaking at 220 rpm in a shaker incubator. The following morning, the cultures were centrifuged for 30 min at 4°C, 3,330 g in an Eppendorf 5810 benchtop centrifuge and the supernatant was discarded. The pellets were resuspended in 5mL fresh medium supplemented with 40% (v/v) glycerol and stored as master stocks at -80°C.

For protein expression, an auto-induction protocol was applied (O'Reilly *et al.*, 2015). Auto-induction was chosen over a conventional IPTG-based expression protocol as it is better suited for simultaneous expression of multiple proteins (Studier, 2005). The auto-induction approach includes glucose and lactose in the medium, requiring the cells to grow to a certain density before protein expression is induced following the eventual depletion of glucose. Clones from master stocks were inoculated into 6 mL 2YT medium supplemented with 2% (w/v) glucose, Ampicillin (100 μg/mL) and Kanamycin (15 μg/mL) overnight at 37^{o}C, being shaken at 220 rpm in a shaker incubator. The following morning the starter culture was refrigerated at 4^{o}C until 5pm. The culture was then inoculated at a 1 in 1,000 dilution into 100 mL of Auto-induction media (ZY medium, 1 mM $MgSO_4$, 1x 5052, 1x NPS, 100 μg/mL Carbenicillin, 15 μg/mL Kanamycin) and grown overnight at 37^{o}C, whilst shaking at 220 rpm. The following morning the cultures were pelleted by centrifugation in an Eppendorf Benchtop centrifuge (3,330 g for 25 min at 4^{o}C) following which the supernatant was removed. The pellets were put on ice and resuspended in Lysis Buffer B (100 mM NaH_2PO_4, 10 mM Tris Cl, 8M Urea, pH 6.3) (5 mL of buffer per gram of pellet) containing protease inhibitors tablets (Roche) at a quarter of the manufacturers recommended concentration (note buffers containing urea were made on the day of use, as high concentrations of urea have a tendency to precipitate out of solutions). Following suspension in Lysis buffer B, the mixture was incubated for lysis for 30 min at 37^{o}C, whilst being shaken at 220rpm. Following incubation, the lysates were returned to ice and sonicated for 15 seconds (2 seconds pulse on, 2 seconds pulse off) at 25% amplitude using a Branson Digital Sonifier (VWR Scientific). The lysate was then centrifuged at 3,330 g for 30 min at 4^{o}C in an Eppendorf benchtop centrifuge. The resultant pellet was discarded and the protein-containing supernatant was passed through a 0.45 μm filter to remove any residual debris.

The His_6-tagged proteins were purified by Immobilised Metal Affinity Chromatography (IMAC) using Ni^+NTA agarose resin on a column. The protocol was carried out at room temperature and based on the 'purification under denaturing conditions' protocol found in the Qiagen's His_6 tagged protein purification handbook (Qiagen, 2003). For purification, 1 mL of the Ni^+ NTA slurry was added to a 5 mL column and the resin was allowed to settle. The column was subsequently equilibrated with 2 mL of Lysis buffer. The filtered bacterial lysate was then passed through the column. This 'flow-through' was collected and passed through the column a second time. The column was then washed with Wash Buffer C (8 M Urea, 100 mM NaH_2PO_4, 10 mM Tris base, pH 6.3), at a volume of three times the initial filtered lysate volume, to remove loosely bound non-specific proteins. The His_6 bound protein-of-interest was then eluted in 6 mL of Elution Buffer B (8 M Urea, 100 mM NaH_2PO_4, 10 mM Tris base, pH 4.5). The eluted proteins were then run on an SDS-PAGE gel for determination of the protein fragment size, and Western Blot (using a HRP-conjugated mouse anti-His_6 antibody (Sigma) diluted 1 in 2,000 in PBST with 1% (w/v) BSA and developed with a TMB substrate) to confirm the presence His_6 tags. Proteins were subsequently quantified using the BCA assay (see Section 2.5.1.3). Detailed SDS-PAGE and Western Blotting protocols can be found in sections 2.3.23 and 2.3.24.

2.5.1.2 Validation of recombinant proteins for screening experiments

Prior to expression and purification of recombinant human proteins obtained from Source Bioscience (Cambridge, UK), each clone was sequenced and analysed using BLAST to confirm it encoded the originally identified protein (Bussow *et al.*, 1998), as the clones were identical to those present on the UniPEx protein arrays. A single colony

of each clone was obtained by generating streak plates from the stocks of each clone. Briefly, the clones were streaked on LB agar plates with carbenicillin (100 μg/mL) and kanamycin (15 μg/mL) and left to incubate overnight at 37^{o}C. Single colonies were picked from each plate and inoculated as stab cultures in a microfuge tube containing 1 mL LB soft-agar supplemented with carbenicillin (100 μg/mL) and kanamycin (15 μg/mL). Clones were incubated overnight at 37^{o}C and then sent to the commercial sequencing service provider GATC Biotech for 'on-site' plasmid extraction and DNA sequencing. The genes were sequenced using the multiple cloning site-flanking primers pQE65 and pQE276 (Büssow *et al.*, 2000) listed below.

pQE65 (sense)

5' - TGA GCG GAT AAC AAT TTC ACA CAG - 3'

pQE276 (reverse)

5' - GGC AAC CGA GCG TTC TGA AC - 3'

Once a sequence was obtained for a clone, a BLAST search was carried out to compare it against NCBI public databases to confirm the identity of the protein. Following BLAST confirmation, the clone sequence was aligned with the NCBI listed sequence using the multiple sequence alignment tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). This illustrated the overlap between the clone and the BLAST-identified protein sequence.

Following confirmation of sequence overlap between the clones and their human proteins, the protein fragments underwent size estimation. Upon confirmation that the clone sequence matched its parent human protein sequence in the NCBI database, the

theoretical size of each protein fragment was estimated using the ExPASy 'Compute pI/MW Tool' (http://www.expasy.org/ tools/pi_tool.html) (Brum *et al.*, 2009). The estimated sizes of the proteins were used in the validation of the successful IMAC purifications of the proteins. Finally, the fragments were also confirmed to be His_6-tagged by probing with peroxidase-conjugated anti-His_6 in western format.

2.5.1.3 Quantitation of proteins using the BCA assay

The BCA Protein Assay (Thermo Scientific™ Pierce™) was used to measure the concentration of protein samples to be used as antigens for screening experiments, instead of a Nanodrop 1000, due to a high urea concentration in the protein's elution buffer. High concentrations of urea can fall out of solutions as crystals, risking interference with Nanodrop 1000 measurements. The BCA assay, however, is suited to quantifying samples containing up to 3M urea. As antigens were purified using 8 M urea, 1 in 10 dilutions were quantified. The BCA assay is based on reduction of Cu^{+2} to Cu^{+1} by proteins in alkaline conditions (the Biuret reaction). This can be measured colorimetrically to detect the level of Cu^{+1} present using a reagent containing bicinchoninic acid (BCA). Each Cu^{+1} molecule chelates two BCA molecules, producing a purple colour that absorbs light at 562 nm. The amount of colour generated increases with protein concentration linearly, up to a concentration of 2,000 μg/mL.

The BCA assays were carried in accordance with the manufacturer's recommended 'Microplate Procedure'. Briefly, BSA standards (0, 25, 125, 250, 500, 750, 1000, 1500 and 2000 μg/mL) were made up (from a manufacturer-provided 2 mg/mL BSA solution) in a 1 in 10 dilution of Elution buffer in PBS (see Section 2.1.7.4 for elution buffer constituents) to ensure the same amount of urea was present in standards as in

samples. Each standard (25 μL) was added in triplicate to wells in a 96-well microplate. The same volume of each antigen, diluted 1 in 10 in PBS, was also added to wells in triplicate. The BCA working reagent was made by mixing 50 parts of BCA Reagent A with 1 part of Reagent B (both manufacturer provided). A total of 200 μL of working reagent was then added to each well, and the plate's contents were mixed on a microplate shaker for 30 sec. The plate was then covered with aluminium foil and incubated away from light in a 37°C incubator for 30 min. Following incubation, the plate was removed from the incubator and allowed to stand for 10 min, to cool to room temperature. The absorbances of the wells were then read at a wavelength of 562 nm using a Tecan Safire2 plate reader. The absorbance values of the BSA standards were used to make a standard curve. This standard curve was then used to determine the protein concentration of each unknown sample.

2.5.2.1 Monoclonal ELISA for screening of 384 avian PCa-enriched scFv against selected antigens identified using a high-density protein array

For screening 384 PCa cell-enriched scFv against purified, protein array-identified antigens, an indirect ELISA was carried out on each scFv in duplicate. Each antigen (PSMA7, RARG, ZNF358 and SETD2) was immobilised in 96 well ELISA plates (Maxisorp™, NUNC™), by incubating 100 μL of each antigen (at a concentration of 10 μg/mL in PBS) in each well for a period of 2 hours at room temperature. The plates were then washed with PBST (x 3) and PBS (x 3), before being blocked with 3% (w/v) BSA in PBS (200 μL per well) overnight at 4°C. The following morning, soluble scFv (see Section 2.3.22 for preparation) were thawed in a 37°C water bath and centrifuged at 3,330 g in

an Eppendorf 5810 benchtop centrifuge at 4°C for 30 min, to pellet debris. The BSA blocking solution was removed from the ELISA plate and the plate was washed with PBST (x 3) and PBS (x 3) as before. The soluble scFv-containing supernatant isolated from each clone (100 μL) was added to the antigen and BSA-blocked well and left to incubate for 1 hour at room temperature. The plate was washed with PBST (x 3) and PBS (x 3) as before and a 1 in 1,000 dilution of a HRP-labelled anti-HA commercial monoclonal antibody (Roche) in 1% (w/v) BSA PBST (100 μL) was added to each well and left to incubate for 1 hour at room temperature. Negative controls were also included on each plate, which excluded antigen, scFv and secondary antibodies from the immunoassay respectively. Each plate was washed with PBST (x 3) and PBS (x 3) and 100 μL of TMB substrate was added per well. Each plate was incubated away from light in a 37°C incubator. The reaction was stopped with the addition of 100 μL 10% (v/v) HCl to each well. The absorbance was then measured at 450 nm using a Tecan Safire 2 plate reader.

2.5.2.2 Investigation of crude lysates for the presence of scFv

For investigating lysates of clones-of-interest for the presence of scFv, 15 μL of crude lysate obtained from deep-well plates (see Section 2.3.22) was mixed 5 μL of 4X Protein loading dye, run on an SDS-PAGE gel, transferred to a nitrocellulose membrane and blocked (see Sections 2.3.23 and 2.3.24). Following blocking, the membrane was washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller mixer at room temperature and subsequently incubated on a roller shaker for 1 hour with a peroxidase-conjugated anti-HA (Roche) diluted 1 in 2,000 in 5 mL of PBST with 1% (w/v) BSA. Following incubation, the membranes were washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller

mixer at room temperature and developed using TMB substrate. The presence of a band of approximately 28 kDa would confirm that the scFv was present in the lysate.

2.5.2.3 Western blots selected scFv-containing lysates of clones-of-interest

Selected scFv-containing lysates of clones-of-interest, identified through monoclonal ELISA screening, were tested against antigens by western blotting. For investigation of lysate-contained scFv binding to antigen, 1 μg of the antigen-of-interest was run on an SDS-PAGE gel (see Section 2.3.23 for protocol) before being transferred to a nitrocellulose membrane for western blot testing (see Section 2.3.24 for protocol). The membranes were left blocking overnight at 4^{o}C under agitation in 5% (w/v) Milk Marvel in PBST. The following morning the membranes were placed in 50 mL polypropelene tubes and washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller mixer at room temperature. Deep-well plates containing the lysate of the scFv were thawed in a water bath at 37^{o}C, and centrifuged at 3,333 g in an Eppendorf 5810 benchtop centrifuge for 30 min at 4^{o}C to ensure the supernatant did not contain debris. Cleared supernatant was diluted 1 in 50 in 5 mL PBST with 1% (w/v) BSA, and was then added to the membrane-containing tubes. The tubes were then incubated on a roller mixer for 1 hour at room temperature. The membranes were washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller mixer at room temperature. Five mL of rabbit anti-HA antibody (Sigma), diluted 1 in 1,000 in PBST with 1% (w/v) BSA, was incubated with the membrane for one hour under the same conditions. The membrane was again washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller mixer at room temperature and then incubated with a peroxidase-conjugated goat anti-rabbit IgG antibody (Sigma) (1 in 5,000) in PBST with 1% (w/v) BSA under the same conditions. The membranes were washed for 5 min in PBST (x 3) and 5 min in PBS (x 3) on a roller mixer at room

temperature for a final time. TMB substrate was added to the membranes to determine whether binding had occurred.

2.5.3.1 Time-point optimisation of P3P2F10

Before proceeding with large-scale scFv expression, the clone P3P2F10 was investigated to determine the optimum IPTG-induction time for expression on a small-scale. For this, a 10 mL starter culture (SB with 100 μg/mL carbenicillin) was inoculated with a scraping from frozen stocks of the clone from the master plates. This culture was set up to incubate overnight at 37°C, whilst shaken at 220 rpm. The following morning the clone was subcultured into 100 mL SB containing 1x 505 and 100 μg/mL carbenicillin at a 1 in 100 dilution. This culture was grown at 37°C, whilst shaking at 220 rpm in a shaker incubator until an OD_{600} nm of 0.6 was reached. IPTG was added to a final concentration of 1 mM for induction of expression, and the temperature of the shaking incubator was changed to 30°C. One mL aliquots were taken from the culture at multiple time-points ranging from 1 hour to overnight incubation for analysis. Each aliquot underwent 3 freeze-thaw cycles in a -20°C freezer and a 37°C microfuge tube heating block to lyse the cells. The lysed cell debris was then pelleted by centrifugation at 10,000 g for 20 min at 4°C in a Hermle microcentrifuge. The supernatant was taken from the cleared lysate from each time point and diluted 1 in 5 in PBST. Diluted lysate (100 μL per well) from each time-point was then added in triplicate to wells of a Nunc Maxisorp 96-well ELISA plate that had been coated earlier with 100 μL of 10 μg/mL ZNF358 in PBS and then blocked with 200 μL of 3% (w/v) BSA in PBS. The diluted lysates were allowed to incubate in the wells for 1 hour at room temperature and the wells were then washed with PBST (x 3) and PBS (x 3). A peroxidase-conjugated anti-HA antibody, diluted 1 in 2,000 in 100

μL of PBST with 1% (w/v) BSA was added to each well. The plate was incubated for one hour at room temperature and then washed with PBST (x 3) and PBS (x 3). TMB (100 μL per well) was then added and allowed to develop for 10 min, following which a difference in colour between time-points was visible. The reaction was subsequently stopped by the addition of 10% (v/v) HCl (100 μL per well) and the absorbance was read at 450 nm on a Tecan Safire 2.

2.5.3.2 Expression and purification of scFv clone P3P2F10 by IMAC

For large-scale expression of the scFv encoded by the clone P3P2F10, a 10 mL starter culture of SB with 100 μg/mL carbenicillin was inoculated with a scraping from frozen stocks of the clone and incubated overnight at 37^{o}C whilst shaking at 220 rpm in a shaking incubator. The following morning the clone was subcultured at a 1 in 100 dilution into two 200 mL cultures of SB with 100μg/mL carbenicillin and 1X 505 (0.5% (v/v) glycerol, 0.05% (v/v) glucose) (in 1 L conical flasks) and grown to an OD of 0.6 at 37^{o}C whilst being shaken at 220 rpm in a shaking incubator. Once the culture had grown to mid-log phase, scFv expression was induced by the addition of IPTG to a final concentration of 1 mM. The culture then was left incubating overnight at 30^{o}C whilst being shaken at 220 rpm in a shaking incubator. The following morning the cultures were pelleted by centrifugation at 3,333 g in an Eppendorf 5810 benchtop centrifuge for 25 min at 4^{o}C. The supernatant was removed and the pellet was resuspended on ice in 10 mL Lysis Buffer B (300 mM NaCl. 50 mM NaH_2PO_4, 10 mM Imidazole, pH 7.5) containing a final concentration of 0.25x Protease inhibitor tablets (Roche). The resuspended pellets were then sonicated on ice for 2 min at 40% amplitude (4 sec pulse on, 2 sec pulse off,) using a microtip Vibra Cell™ sonicator. The cell lysate was then pelleted by

centrifugation at 15,000 g for 25 min in an Eppendorf 5810 benchtop centrifuge at 4^{o}C. The supernatant was then taken and filtered at 0.2 μm to remove any non-pelleted cell debris.

The filtered scFv-containing lysate then underwent IMAC purification at 4^{o}C. Amintra Resin slurry (2 mL) was added to a 10 mL column and left until the Ni^{+}-NTA resin had settled. At no stage during this process was the Ni^{+}-NTA resin allowed to dry. The column was then equilibrated by passing 4 mL of Lysis Buffer B through it. The lysate was then passed through the column and collected. The flow-through was passed through the column again to maximise the amount of His_6-tagged scFv bound to the column. For removal of non-specifically bound proteins, the column was then washed once with 8 mL Wash Buffer A (50 mM NaH_2PO_4, 1 M NaCl, 10% (v/v) glycerol, 20 mM Imidazole, 1% (v/v) Triton X-100, pH 7.5) and then twice with 8 mL Wash Buffer B (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazole, pH 7.5). Samples of the filtered lysate flow-through, as well as samples from both wash fractions A and B, were taken for analysis. For elution of resin-bound protein, 10 mL Elution Buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM Imidazole, pH 7.5) was passed through the column and collected as 10 x 1 mL fractions. The fractions were to be used for SDS-PAGE and Western Blot analysis. The protein concentration of each fraction was checked using a Nanodrop™ 1000 and those fractions containing protein were pooled and buffer exchanged against 12 mL of 0.2 μm filter-sterilised PBS containing dissolved protease inhibitor tablets (Roche) at the manufacturer's recommended working concentration. For the buffer exchange, two 5 kDa molecular weight 'cut-off' VivaspinTM 6 columns were used. The columns were loaded with 6 mL of eluted samples and centrifuged at 3,330 g in an Eppendorf 5810 benchtop centrifuge at 4^{o}C until approximately 5.5 mL had passed through. At this point

the flow through was discarded and column was reloaded with remaining solution. The centrifugation step was repeated until 0.5-1 mL remained. The remaining protein solution was then aliquoted and stored at -20^{o}C.

2.5.3.3 Indirect ELISA characterisation of the scFv clone P3P2F10 binding to ZNF358

Checkerboard ELISA analysis was carried out to investigate the binding of the scFv clone P3P2F10 to ZNF358 at varying concentrations of both antibody and antigen. For this, the antigen was coated at varying concentrations in triplicate (100 μL volume per well) on 96-well Nunc Maxisorb™ plates for two hours at room temperature. The plates were washed with PBST (x 3) and PBS (x 3). Following washing, each well was blocked with 200 μL PBS with 3% (w/v) in BSA and the plates were left to incubate overnight at 4^{o}C. The following morning the plates were washed with PBST (x 3) and PBS (x 3) and different dilutions of the scFv in PBST with 1% (w/v) BSA (100 μL per well) were added to each well to incubate for 1 hour at room temperature. The plates were washed with PBST (x 3) and PBS (x 3) and 100 μL of a rabbit anti-HA (Sigma), diluted 1 in 1,000 in PBST with 1% (w/v) BSA, was added to each well and left to incubate for 1 hour at room temperature. The wells were again washed with PBST (x 3) and PBS (x 3) and 100 μL of a peroxidase-conjugated goat anti-rabbit IgG antibody (Sigma) diluted 1 in 5,000 in PBST with 1% (w/v) BSA was added to each well and left to incubate for 1 hour at room temperature. Following a final wash with PBST (x 3) and PBS (x 3), the wells were developed with the addition of 100 μL TMB for ten minutes before the reaction was stopped with the addition of 10% (v/v) HCl. The absorbances were then read at 450 nm using a Tecan Safire2.

2.5.3.4 Antigen cross-reactivity of the scFv clone P3P2F10

The scFv P3P2F10 was investigated in an indirect ELISA format for cross-reactivity with three proteins, identified as positive antigens through UniPEx Protein Array Experiments (RARG, SETD2 and ZNF358), one protein identified through UniPEx Protein Array Experiments (RAS dmi) and one protein not present on the array (a non-specific scAb developed 'in-house'). Antigens (100 μL) were added at a concentration of 10 μg/mL on 96-well Nunc Maxisorb™ Plates and left to incubate for two hours at room temperature. The plates were washed with PBST (x 3) and PBS (x 3) and 200 μL of PBST with 3% (w/v) BSA was added to wells to block overnight at 4°C. Following blocking the wells were washed with PBST (x 3) and PBS (x 3). The 100μL of the P3P2F10 scFv, diluted 1 in 1,000 in PBST with 1% (w/v) BSA, was added to each well and left to incubate for 1 hour at room temperature. The wells were then washed with PBST (x 3) and PBS (x 3) and 100 μL of a commercial HRP-conjugated anti-HA antibody (Roche), diluted 1 in 1,000 in PBST with 1% (w/v) BSA, was added to each well. The plates were left to incubate at room temperature for 1 hour and then washed with PBST (x 3) and PBS (x 3). The wells were developed with the addition of 100 μL TMB for ten minutes before the reaction was stopped with the addition of 100 μL 10% (v/v) HCl. The absorbances were then read at 450 nm using a Tecan Safire2 plate reader.

2.5.3.6 Western blot analysis of the scFv clone P3P2F10 binding to ZNF358

For analysis of scFv clone P3P2F10 binding to ZNF358 in western blotting format, 1 μg of the antigen ZNF358 was run on an SDS-PAGE gel and transferred to a nitrocellulose membrane (protocol detailed in 2.3.24). The antigen-bound membrane was blocked in

PBST with 5% (w/v) MM overnight. The membrane was then washed for 5 min in PBST (x 3) and 5 min PBS (x 3) on a platform shaker. Following washing, the membrane was incubated with 5 ml of the P3P2P10 scFv (stock concentration 0.25 mg/mL) at a 1 in 1,000 dilution in PBST with 1% (w/v) BSA and left to incubate for 1 hour at room temperature on a roller shaker. The membrane was then washed for 5 min in PBST (x 3) and 5 min PBS (x 3) on a platform shaker. For detection of bound scFv, a commercial rabbit anti-HA antibody diluted 1 in 1,000 in 5 mL of PBST with 1% (w/v) BSA was added and left to incubate for 1 hour on a roller shaker. The membrane was then washed for 5 min in 20 mL PBST (x 3) and 5 min in 20 mL PBS (x 3) on a platform shaker, followed by incubation with a commercial peroxidase-conjugated goat anti-rabbit IgG (Sigma) diluted 1 in 5,000 in 5 mL PBST with 1% (w/v) BSA and incubated for 1 hour on a roller shaker at room temperature. The membrane was washed for 5 min in PBST (x 3) and 5 min PBS (x 3) on a platform shaker for a final time and the substrate TMB was added to the membrane to detect the presence of bound antibody.

2.5.3.7 Characterisation of P3P2F10 anti-ZNF358 binding with a Biacore 3000

For 'label-free' binding analysis of the scFv P3P2F10 to the antigen ZNF358, a Biacore 3000 instrument was used. The Biacore 3000, utilizes surface plasmon resonance (SPR) to measure biomolecular binding interactions on its sensor chip surface without the need for labels. The chip surface is made up of a dextran-layer (for protein immobilisation) that is covalently attached to a prism-linked gold layer (to allow for evanescent wave generation during surface plasmon resonance). The instrument operates on the premise that, when a molecule is immobilised to its chip surface, the

angle at which SPR is achieved will change. When an analyte is passed over the surface and binds to the immobilised molecule, another change in the resonance angle will be induced. The platform monitors the changes in the resonance angle in real-time, allowing all binding events to be recorded. In this protocol, the antigen ZNF358 was immoblised to the surface and the scFv was passed over it in solution.

Prior to the experiment, the antigen underwent buffer exchange into degassed, sterile filtered (using a 0.2 μm pore filter) HBS-EP$^+$ buffer (0.1 M HEPES, 1.5 M NaCl, 30 mM EDTA, 0.05% (v/v) Tween 20 in MilliQ H_2O, pH 7.4). For this, 500 μL of the antigen was added to a 5 kDa molecular weight 'cut-off' VivaspinTM 6 column and was buffer exchanged with 20mL sterile filtered and degassed HBS-EP$^+$ buffer, by centrifugation in an Eppendorf 5810 benchtop centrifuge at 4^{o}C, to yield a stock of 0.18 mg/mL ZNF358 for Biacore work (determined by Nanodrop 1000^{TM}).

With a maintenance chip docked, the Biacore^{TM} 3000 instrument underwent cleaning using a 'sanitisation' and 'desorb' program at 25^{o}C to remove residual contaminants before kinetic binding analysis. The maintenance chip was replaced with a research-grade carboxymethlyated (CM) dextran sensor chip (CM5 sensor chip) (GE Healthcare). The dextran-layer of the chip is covalently attached to a gold layer capable of undergoing surface plasmon resonance. The dextran surface was activated by mixing equal volumes (70 μL each) of 400 mM ethyl(dimethylaminopropyl)-carbodiimide (EDC) and 100 mM N-hydroxysulfosuccinimide (NHS) and passing 80 μL of the mixture over the sensor surface at a flow rate of 10 μL/minute. The antigen (ZNF358) was immobilised at a 10 μg/mL concentration in 10 mM NaOAc, pH 3.9, onto the chip surface of flow cell 4 (FC4) at a flow rate of 10 μL/min for 20 minutes. The surface was then capped by passing 1M

ethanolamine HCl, pH 8.5, over the surface for 7 minutes. The surface was further cleaned with four 30 second injections of 10 mM NaOH at a flow rate 10 μL/minute, removing any unbound or extraneous material. A binding response of 275 resonance units (RU) was obtained, indicating successful antigen binding. Flow cell 3 (FC3) was used as a reference (negative control). This was activated and capped in the same manner as FC4 but without antigen immobilisation. The response of the reference chip was subtracted, in real-time, from that of antigen-bound chip, to indicate the true antibody-antigen binding kinetics.

The scFv P2P3F10 at concentrations of 8.93 nM (0.25 μg/mL), 2.23 nM (0.06 μg/mL) and 0.89 nM (0.02 μg/mL) in sterile filtered and degassed HBS-EP$^+$ buffer was passed over the ZNF358-immobilised and reference surfaces (FC4 and FC3) at a flow rate of 10 μL per minute and binding levels were determined in real-time. The surface was regenerated using 10 μL 10 mM NaOH. The 2.23 nM concentration of antibody was run in duplicate and a zero concentration (buffer-only control) was also included, to enable double referencing. Association and dissociation phases were monitored for 4.5 and 10 min, respectively. Sensograms for each of the scFv concentrations were fitted globally to a 1:1 interaction model using Biacore 3000 dedicated software.

Chapter 3 – Generation of a recombinant antibody library against prostate cancer cells

3.1 Introduction

Prostate Cancer (PCa) is the most commonly diagnosed form of cancer in men in developed countries (Heidenreich *et al.*, 2014a). Current methods to diagnose PCa are limited to the DRE and PSA assay, which lack sensitivity and specificity (Gilgunn *et al.*, 2013). More reliable diagnostic approaches, such as the biopsy, are too invasive to be used in routine testing. Treatment options are limited, and in the case of advanced PCa, focus on prolonging life rather than curing the disease (Heidenreich *et al.*, 2014b). Thus, new approaches to the diagnosis and treatment of PCa are greatly needed. Antibodies are a well-established means, to diagnose and to treat cancer. The discovery of novel molecular targets in cancer, and the generation of antibodies against them, would bring huge advancements to the field (Popkov *et al.*, 2004). However, the most established methods of generating antibodies rely on selecting an antigen before the process, limiting its application in target discovery (Conroy *et al.*, 2012).

Cell-based animal immunisation, coupled with cell panning using phage display technology, has the capacity to generate a diverse range of antibodies against both known and unknown cancer-associated antigens (Hoogenboom, 2005; Even-Desrumeaux *et al.*, 2013). However, this non-targeted approach has significant difficulties in isolating antibodies of value, and ultimately relies on extensive characterisation experiments, followed by antigen identification through MS (Keller *et al.*, 2015). Hence, improved approaches, to develop and characterise cancer-specific antibodies, are badly needed. This research aims to address these needs. The goal of the work described in this chapter was to produce a panel of antibodies, enriched against antigens present in prostate cancer cells, that were not predetermined. This antibody

panel was required as a prerequisite to developing and demonstrating an alternative approach to antibody discovery.

Chickens were chosen for immunisations instead of other standard hosts, such as mice or rabbits, due to their increased phylogenetic distance to the immunogen (human prostate cancer cells). The rationale for this approach, was that there should be a maximal immune response to the antigens present in PCa cells, based on their reduced homology with the native proteins already present in the avian host. Furthermore, the immune system of chickens, unlike mammals, has single functional variable genes for both the heavy and light chains of their immunoglobulins (Wu *et al.*, 2012). Therefore, fewer experimental steps are required to convert an avian immune response into a recombinant scFv library. For these reasons, chickens have been shown to be ideal hosts for antibody generation against human proteins in the past (Leonard *et al.*, 2007). Two chickens were immunised with differing combinations of PCa cell lines to elicit an anti-PCa immune response. Upon completion of the immunisation regime, both immune responses underwent conversion into single chain Fv (scFv) pools and were cloned into the pComb3X phagemid vector, creating two antibody libraries. One library was selected, displayed as scFv-phage and subsequently underwent multiple rounds of panning against PC-3 cells. Following this anti-PCa enrichment strategy, the scFv encoded by the enriched scFv-phage were transferred into a non-suppressor strain for soluble expression and isolation.

3.2 Results

3.2.1 Immunisation of two chickens with PCa cell lines

Female leghorn chickens (ABG055 and ABG048) underwent an immunisation regime against a combination of PCa cell lines (see methods section 2.3.2). After the second boost, the antibody serum titre was determined, to ensure PCa-specific antibodies were being generated (see methods section 2.3.3). Both chickens, ABG055 and ABG048, exhibited a significant immune response to the cells (see Figures 3.1 and 3.2). In both experiments, sera yielded similar absorbances at a dilution of 1 in 100. This suggests non-specific binding, and can be explained as the Hook effect - which is a reduction in binding in more concentrated serum samples. Serum contains a multitude of potential immunoassay-interfering substances, such as heterophilic antibodies and complement proteins, the effect of which can be nullified through increasing the dilution factor (Tate and Ward, 2004). Higher dilutions of serum isolated from both chickens exhibited an increase in binding signal, supporting this. In both experiments it was observed that the IgY-binding signal was higher in serum isolated at later stages in the immunisation regime. Serum isolated following the second boost of both chickens yielded a binding signal higher than the control at dilutions as high as 1 in 1,000,000, which is more than a sufficient immune response for antibody generation.

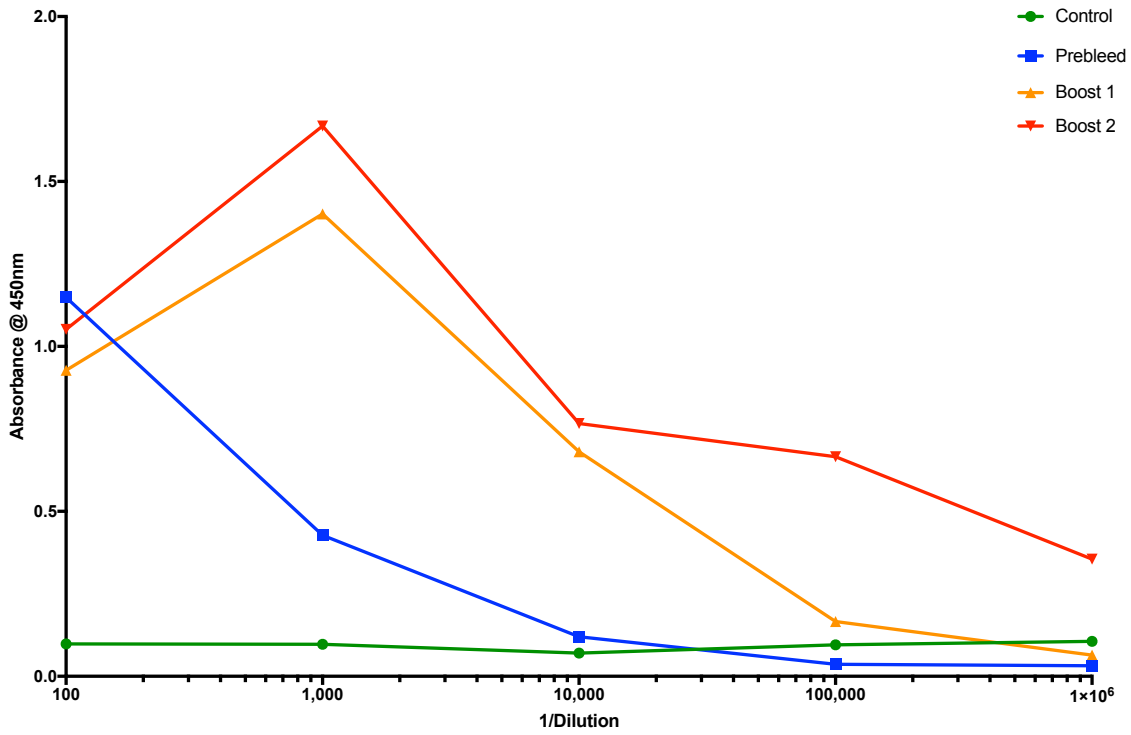

Figure 3.1 – Antibody serum titre of the Chicken ABG055 against PC-3 cells by indirect ELISA. The chicken was immunised over a 4-month period, with boosts 6 and 13 weeks after the primary immunisation. Serum samples were taken before an initial immunisation with PCa cells (Prebleed), one week after the first boost (Boost 1) of PCa cells was administered, and one week after the second boost (Boost 2) of PCa cells was administered (Boost 2). PBS was used as a negative control (Control). Bound antibodies were detected by HRP-labelled rabbit anti-chicken followed by the addition of the substrate, TMB. The absorbance was read at 450 nm.

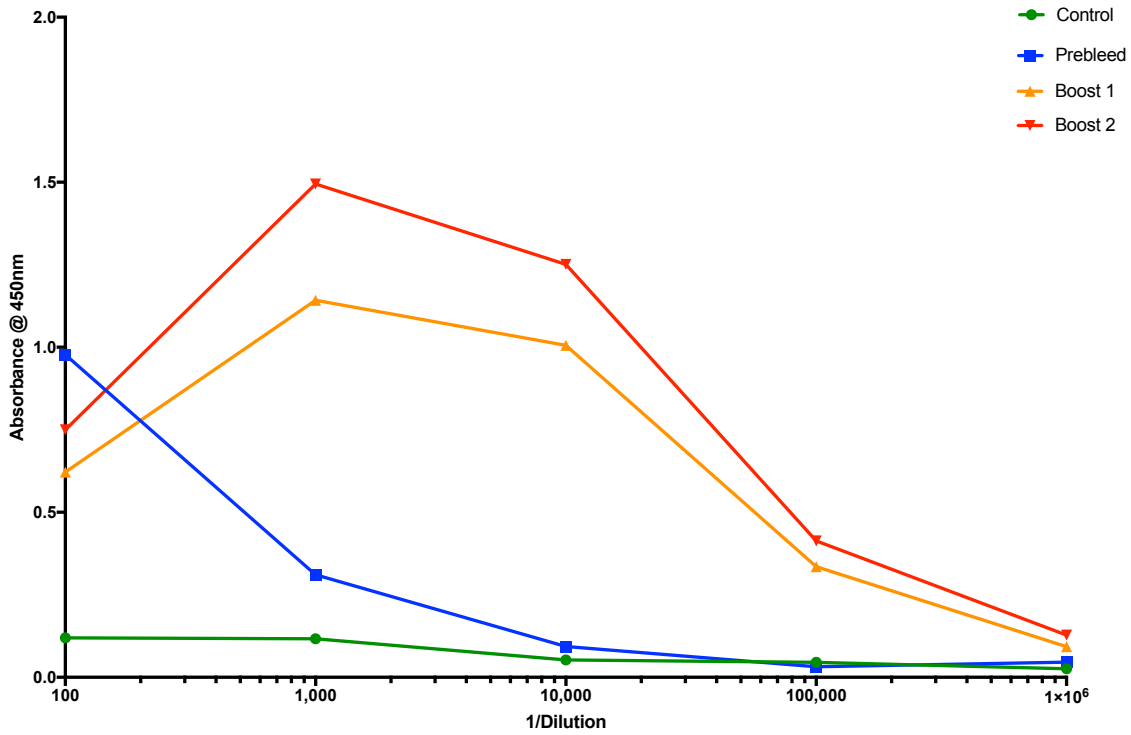

Figure 3.2 – Antibody serum titre of the Chicken ABG048 against PC-3 cells by indirect ELISA. The chicken was immunised over a 4-month period, with boosts 6 and 13 weeks after the primary immunisation. Serum samples were taken before an initial immunisation with PCa cells (Prebleed), one week after the first boost (Boost 1) of PCa cells was administered, and one week after the second boost (Boost 2) of PCa cells was administered (Boost 2). PBS was used as a negative control (Control). Bound antibodies were detected by HRP-labelled rabbit anti-chicken followed by the addition of the substrate, TMB. The absorbance was read at 450 nm.

3.2.2 Extraction of RNA from chicken spleens

Following confirmation of an immune response, the chickens (ABG055 and ABG048) were given a final boost and sacrificed seven days later. Their spleens were removed, from which the total mRNA was isolated (see methods section 2.3.4). The mRNA from both chickens was run on an agarose gel and investigated for degradation (see Figure 3.3). Intact 18S and 28S ribosomal RNA bands were observed, indicating that the isolated RNA was of good quality. The concentration of both RNA pools was quantified using a Nanodrop 1000 and deemed sufficient to proceed with cDNA synthesis by reverse transcription (100 μg of RNA was required for cDNA synthesis reactions) (see materials and methods section 2.3.5). Both RNA and cDNA yields are displayed in Table 3.1.

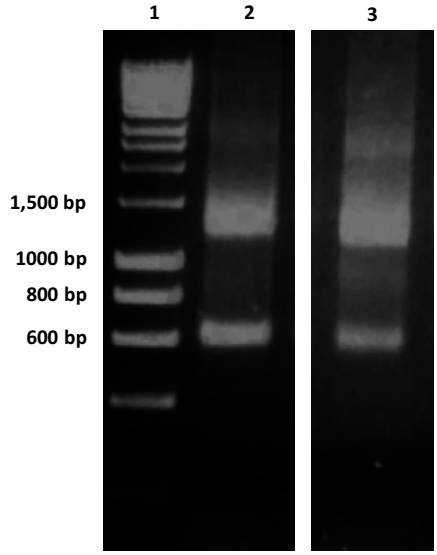

Figure 3.3 – RNA isolated from the spleens of chickens ABG055 and ABG048 separated by gel electrophoresis. Lane 1 contains DNA Hyper Ladder, used as a size reference. Lane 2 contains RNA isolated from the chicken ABG048. Lane 3 contains RNA isolated from the chicken AGB055. Intact 28S and 18S ribosomal RNA bands are visible at ~600 bp and ~1.5 kbp.

Table 3.1 – RNA and cDNA yields following spleen RNA extraction and cDNA synthesis protocols. Note – RNA extract from spleens had a total volume 500 μL. Total volume for cDNA was 560 μL.

Yield	ABG048	ABG055
RNA (ng/μL)	3,525	5,716
cDNA (ng/μL)	1,541	1543

3.2.3 Amplification of variable heavy and light genes and subsequent purification

Following cDNA synthesis, antibody library construction was carried out based on methods described by Andris-Widhopf *et. al* (2000). The variable heavy and light genes from the cDNA pools from both chickens (ABG055 and ABG048) were amplified using two separate PCR reactions. The PCR products were subsequently run on an agarose gel to assess whether sufficient amplification had occurred (see Section 2.3.6 for protocol). Successful variable heavy and light chain amplification was achieved from both cDNA sources and can be seen in Figures 3.4 and 3.5. To generate sufficient quantities of each fragment for scFv gene generation via SOE-PCR, large-scale (10x) reactions were carried out. The resulting products were cut from an agarose gel after electrophoresis and were purified (see methods section 2.3.7).

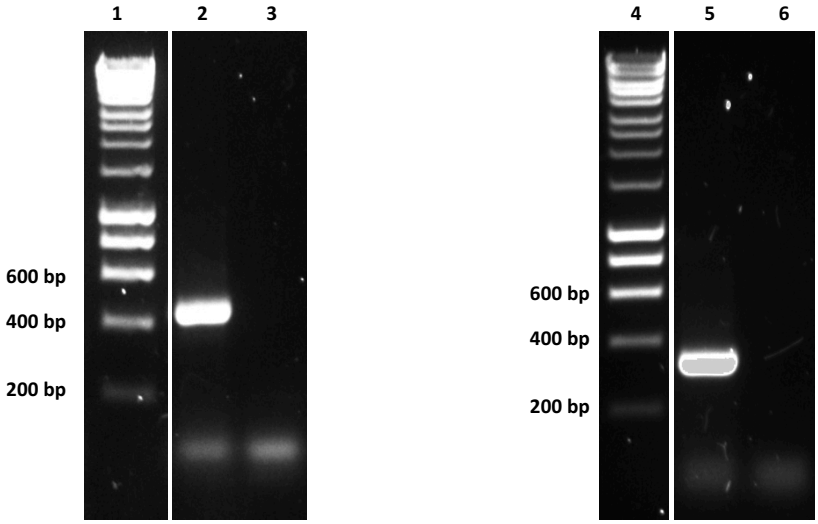

Figure 3.4 – Chicken ABG055 V_H and V_L regions amplified from cDNA and run on a 1.5% (w/v) agarose gel by electrophoresis. Lanes 1 and 4 contain DNA Hyper Ladder to provide a size reference for the amplicons. Lane 2 contains the amplicon from the V_H gene amplification, providing a product of the expected size (~400 bp). Lane 3 contains amplicons from a negative control, using the same constituents used to generate the product in lane 2 excluding template DNA. Lane 5 contains the amplicon from the V_L reaction and is also of the expected size (~340 bp). Lane 6 contains amplicons from a negative control, using the same constituents used to generate the product in lane 5 excluding template DNA.

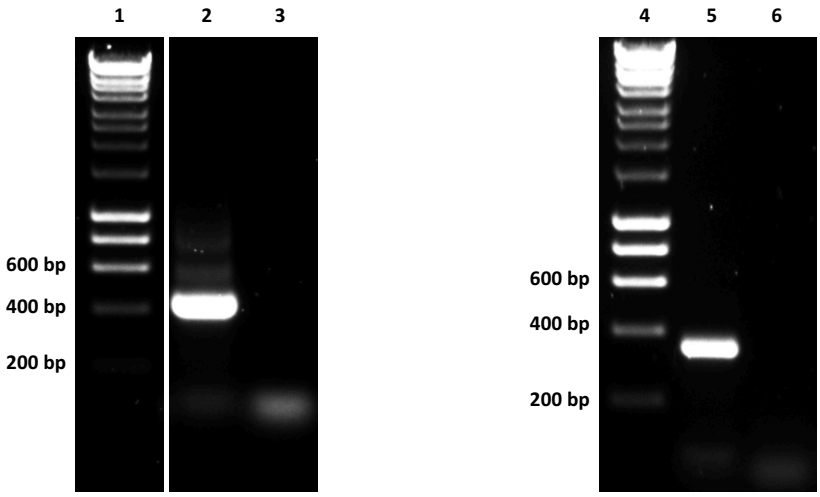

Figure 3.5 – Chicken ABG048 V_H and V_L regions amplified from cDNA and run on a 1.5% (w/v) agarose gel by electrophoresis. Lanes 1 and 4 contain DNA Hyper Ladder to provide a size reference for the amplicons. Lane 2 contains the amplicon from the V_H gene amplification, providing a product of the expected size (~400 bp). Lane 3 contains amplicons from a negative control, using the same constituents used to generate the product in lane 2, excluding template DNA. Lane 5 contains the amplicon from the V_L reaction and is also of the expected size (~340 bp). Lane 6 contains amplicons from a negative control, using the same constituents used to generate the product in lane 5, excluding template DNA.

3.2.4 Generation of scFv fragments by SOE PCR

Following the successful generation of V_H and V_L gene pools from the cDNA of the chickens ABG048 and ABG055, the amplified products were joined together with a serine-glycine linker (GGSSRSSSSGGGGSGGGG) using the sequence tails introduced in the previous rounds of PCR. The scFv product from the reaction was expected to be ~750 bp. This amplification was achieved by SOE-PCR using the flanking CSC-F sense and CSC-B antisense primers (see methods section 2.3.8). These primers introduced *Sfi*I restriction sites at either side of the scFv gene, to allow for unidirectional cloning into a phagemid vector. The expected ~750bp products were identified from both ABG055 and ABG048 reactions (see Figure 3.6), in addition to non-specific products, which were subsequently removed through gel purification. The SOE fragments generated via PCR were produced in larger quantities (10µg) than the minimum required for transformation (700ng) to allow for potential loss of product following proceeding enzymatic treatments and ethanol precipitations. Purified SOE fragments were pooled and stored at -20^{o}C until used in vector cloning experiments.

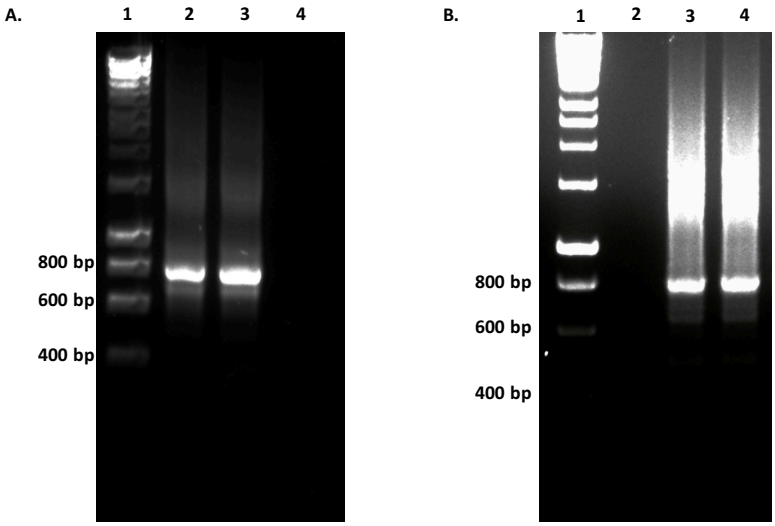

Figure 3.6 - Splice by Overlap Extension PCR of the variable heavy and light genes, generated from the cDNA pools of the chickens ABG048 (Gel A) and ABG055 (Gel B), separated by electrophoresis on a 1% (w/v) agarose gel. Hyper Ladder is present in lane 1 of both A and B gels to allow for size estimation of the products. Both gels display specific bands of ~750 bp in test lanes (lanes 2A, 3A, 3B and 4B), indicating successful fusion of V_H and V_L chains. In addition, the negative control (template DNA excluded from the reaction) yielded no product (lanes 4A and 2B), as expected.

3.2.5 Preparation of a pComb3XSS phagemid vector for SOE cloning

Like the scFv gene pool, the ~3.3kb pComb3XSS phagemid vector to be used in library generation contained two *Sfi*I restriction sites for directional cloning. These sites flanked a removable stuffer fragment (SS) present in the vector (~1.6kb in size). Large quantities of the pComb3XSS vector underwent *Sfi*I restriction. Following restriction, the products were run on a 1% (w/v) agarose gel (see Figure 3.7), for 3.3kb phagemid vector fragment purification (see Section 2.3.10). The SOE gene pools, previously generated from chicken ABG055 and ABG048 cDNA, were also *Sfi*I restricted and gel-purified. Following isolation of the 3.3 kb vector fragment, the 5' phosphate group was removed by Antarctic phosphatase treatment (see Section 2.3.11). Phosphatase-treated fragments lack the 5' phosphoryl terminal required for ligation and, hence, cannot self-ligate when treated with ligase.

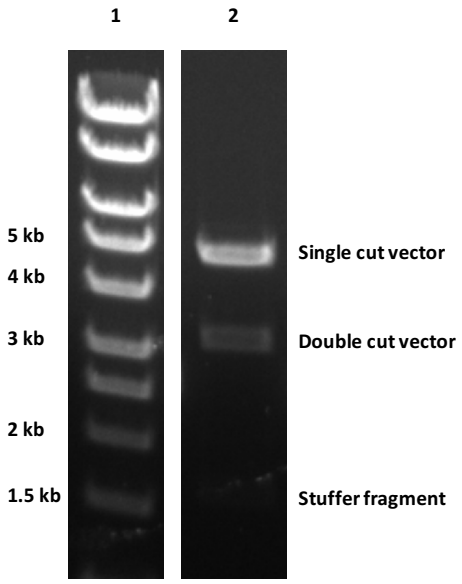

Figure 3.7 - Restriction digest of a pComb3XSS vector with the enzyme *Sfi*I. Following restriction, the DNA was run on a 1% (w/v) agarose gel for separation of pComb3X vector from pComb3XSS vector and stuffer fragment. Lane 1 contains Hyper ladder for a size reference to identify DNA products. Lane 2 contains *Sfi*I-treated vector DNA. DNA bands in the lane are: Stuffer fragment (1.6kb band), double-cut vector (3.3kb band) and an *Sfi*I single-cut vector (5kb band).

3.2.6 Cloning of scFv library into the pComb3X vector and library size estimation

Following enzymatic preparations of pComb3X vector and ABG055/ABG048 scFv gene pools, the scFv pools were ligated into the vector, creating phagemid DNA libraries (see Section 2.3.15) (see Figure 2.1 for schematic). Both DNA libraries underwent ethanol precipitation (see Section 2.3.12) to remove residual enzymes and salts. The phagemid libraries then underwent transformation into high efficiency electro-competent *E. coli* XL-1 Blue cells (see methods section 2.3.16). The size of each library was subsequently calculated (see Table 3.2). While both libraries were calculated to be of a suitable size for panning experiments, it is preferable to pan a library of a size >5 x 10^7 (Barbas III *et al.*, 2001). Therefore, the ABG055 library was more likely to contain scFv of interest. Hence the ABG055 library was selected for enrichment through phage-display.

Table 3.2 - Calculated size of scFv libraries.

Library	Size (cfu/mL)
ABG048	1.8 x 10^6
ABG055	6 x 10^8

3.2.7 Investigation of the ABG055 scFv library diversity through sequencing

Prior to panning, five individual clones from the ABG055 library were sequenced and compared, to ensure diversity, using *ompseq* and *gback* primers (see Sections 2.3.17 and 2.3.18). The nucleotide sequences obtained were translated into amino acid sequences using the Expasy (Expert Protein Analysis System) translate tool, and the Complementarity Determining Regions (CDRs) were identified (see Table 3.3 for light chain CDRs and 3.4 for heavy chain CDRs) based on the Kabat numbering system (Martin, 2015). Analysis of the CDR regions showed that no two clones shared the same scFv gene; hence, the library was considered suitable to undergo panning.

Table 3.3 - Amino Acid sequences of the Light Chain Complementarity Determining Regions from 5 clones isolated from the ABG055 library.

Clone	CDR-L1	CDR-L2	CDR-L3
1	SGGGNYYG	YNTNRPS	GNADSSGTGI
2	SGGSYSYG	WSDKRPS	GSEESNGDGI
3	SGSSGSYG	SNDKRPS	GSWESYVDNGFPGI
4	SGSSGYGYG	SNDKRPS	GSASSSNTYADI
5	SGGSSSYYG	ENNKRPS	GSWEDSSYFGI

Table 3.4 - Amino Acid sequences of the Heavy Chain Complementarity Determining Regions from 5 clones isolated from the ABG055 library.

Clone	CDR-H1	CDR-H2	CDR-H3
1	GFSFSDYNMA	IRKDGGDTHYGAAVKG	TAYGCTWGGCANIIDS
2	GFTFSSFNMF	IENDGSSTRYGPAVEG	SNYGGYWSCPYCIDEIDA
3	GFDLSSYHIH	SAYTGDYTYTAAAVKG	AADDCSWGGCVDNIDA
4	GFTFSSHSMH	ISGDASITRYAPAVKG	ESGSGGSGYCDSSTGCIDS
5	GFDFSSYQMN	INRFGNSTGHGAAVKG	HVYGYCGRGGWCAGLGEIDA

3.2.8 Expression of scFv-phage library and enrichment against PCa cells by direct cell panning

The ABG055 scFv-phage library underwent four rounds of live-cell panning against PC-3 cells, based on the protocol described by Barbas III *et al.* (2001). The aim of this was to introduce a selective pressure, to isolate antibodies specific to PCa cell-associated proteins from the overall library. An initial enrichment strategy was applied, in which the scFv-phage were subjected to an increased number of washes, while exposed to a decreased number of PC-3 cells, with each round of panning. However, this approach was determined to have been unsuccessful, after investigation by polyclonal scFv-phage ELISA (data not shown). The panning strategy was then revised (see Table 2.2); increasing the wash steps between rounds, while keeping the cell number constant (see Section 2.3.19).

3.2.9 Investigation of panning success via polyclonal phage ELISA and 'colony pick' PCR

The precipitated scFv-phage from each round of panning was incorporated into a polyclonal scFv-phage ELISA (see Figure 3.8), to investigate for enrichment to PC-3 cells between rounds (see Section 2.3.21). Enrichment against the cells was achieved, particularly after round 3, which yielded a binding signal >10x higher than that of the unpanned library. A subsequent reduction in signal in Pan 4 was also observed, likely due to the increase in panning stringency of the round.

The panning process was also monitored at different stages by 'colony pick' PCR for the continued presence of scFv fragments (see methods 2.3.20). Contamination of output phage, by phage lacking an scFv gene, risks undesirable clones propagating at a faster rate than 'scFv-harbouring' clones. This would negatively impact the enrichment strategy. Thus, PCR amplicons, of five single colonies from rounds 0, 2 and 4 (see Figure 3.9) were run on a 1% (w/v) agarose gel to investigate for the presence of scFv genes. It was observed that all clones analysed contained a scFv gene of ~750 bp in size. An additional fragment of ~450 kb was also seen in two of the clones tested in the unpanned library. This was partial amplification of the full fragment, which can occur in the presence of repetitive sequences (Hommelsheim *et al.*, 2014).

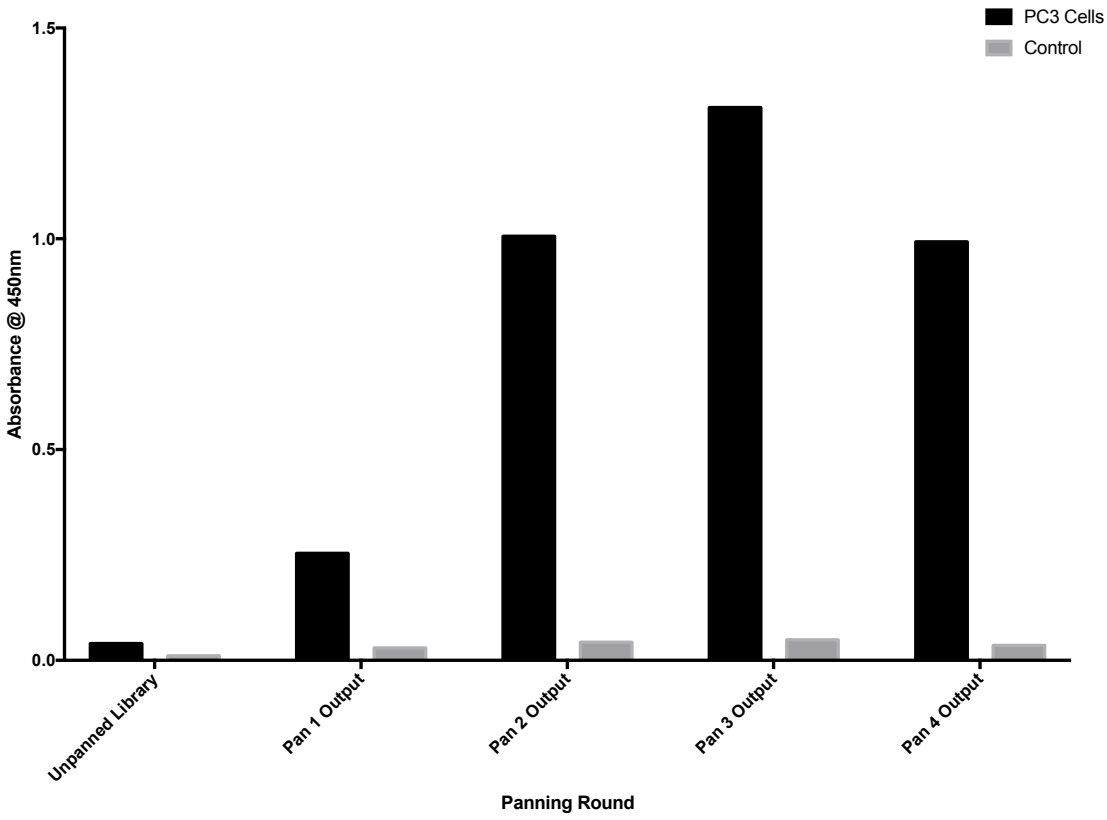

Figure 3.8 - Polyclonal indirect ELISA of output scFv-phage against PC3 cells following each round of panning (black columns). Each phage pool was also tested against wells incubated with FBS-supplemented RPMI media before being blocked with 5% (w/v) Milk Marvel as a control (grey columns). Bound phage were detected using a HRP-conjugated anti-M13 antibody and the absorbance was read at 450nm.

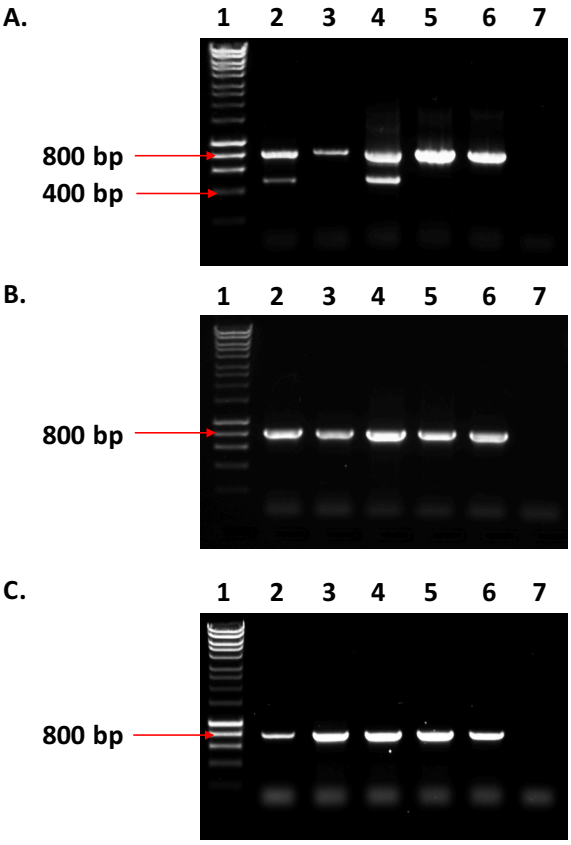

Figure 3.9 - Colony pick PCRs on randomly selected clones from panning rounds 0 (A), 2(B) and 4(C) to monitor for the presence of scFv genes throughout the panning process. The layout of all gels were identical – Lane 1 contains a 1kb Hyper Ladder, Lanes 2-6 contain amplicons from randomly picked single colonies, Lane 7 contains all reaction constituents minus the bacterial colony as a control. Successfully amplified scFv gene inserts are visible as bands of ~750 kb in size.

3.2.10 Isolation of 384 monoclonal soluble scFv enriched against PCa cells

Output scFv-phage pools from panning rounds three and four were considered more likely to contain binders to antigens present in PCa-cells. Before screening for monoclonal scFv-of-interest in the scFv-phage pools, the scFv were expressed in the absence of the phage particle, as bacteriophage have the tendency to bind non-specifically to surfaces, potentially interfering with downstream characterisation experiments (Jones *et al.*, 2016). ScFv isolated from Round 3 were prioritised over those from Round 4, due to the higher binding signal observed in the polyclonal phage ELISA. The clones were sub-cultured into deep-well plates and scFv expression was induced (see Section 2.3.22). These plates were stored at -80^{o}C with protease inhibitors, until required for protein array-based characterisation and screening experiments (chapters 4 and 5).

3.2.11 Investigation of diversity of scFv enriched against PCa cells

Before proceeding with scFv characterisation experiments in chapter 4, five clones from Round 4 output plates (master plate clones C11, G1, H5, F11 and G7) were sequenced using *ompseq* and *gback* primers (see Section 2.3.17). The nucleotide sequences were translated to AAs and the CDRs were compared for diversity using the Kabat numbering system (Martin, 2015). Three unique scFv sequences were obtained from the 5 clones (see Tables 3.5 and 3.6). Three clones (F11, G1 and G7) were identical, and therefore, only listed once in the table as 'Clone 2'. All clones shared the same CDRH2 and 3 domains. These CDR sequences were shorter than those analysed from the unpanned library (see Table 3.4), potentially granting them a higher rate-of-propagation. The shared CDR sequences are evidence of selective pressure, applied through the panning experiments. The variable light chain regions were unique. Diversity in one CDR region alone is sufficient for selective binding between multiple antigens (Xu and Davis, 2000). The location of such diversity is flexible to both the heavy and light chains (Persson *et al.*, 2013).

Table 3.5 - Amino Acid sequences of the Light Chain Complementarity Determining Regions from 5 clones enriched against PCa cells through 4 rounds of panning. The CDR sequences from the fourth and fifth clones were identical to clone 2 and hence excluded from the table.

Clone	CDR-L1	CDR-L2	CDR-L3
1	SGGSGSYG	WNDKRPS	GNADSSGTGI
2	SGGGSRSYYG	NNNKRPS	GGYDSSTDAGI
3	SGGSNNYAGSYYYG	SNDKRPS	GSGDSSNTAGI

Table 3.6 - Amino Acid sequences of the Heavy Chain Complementarity Determining Regions from 5 clones enriched against PCa cells through 4 rounds of panning. The CDR sequences from the fourth and fifth clones were identical to clone 2 and hence excluded from the table.

Clone	CDR-H1	CDR-H2	CDR-H3
1	KAFGFTFSDRGMG	IRSDGSVTNYGSAVKG	DVVVTVLGVFAGHIDA
2	GFTFSDRGMG	IRSDGSVTNYGSAVKG	DVVVTVLGVFAGHIDA
3	GFTFSDRGMG	IRSDGSVTNYGSAVKG	DVVVTVLGVFAGHIDA

3.3 Discussion

This chapter describes the generation and isolation of 384 scFv against PCa cell antigens, to be subsequently incorporated into a novel antigen characterisation experiment.

Here, chickens were immunised with differing panels of PCa cell. One chicken, ABG048, underwent an immunisation regime using four PCa cell lines, to generate a response against antigens present in multiple phenotypes of PCa. The other, ABG055, was immunised with two androgen-independent lines to generate a more focused response against the more aggressive form of the disease (Thalmann *et al.*, 1994; Nemeth *et al.*, 1999). Chickens are excellent hosts for the generation of specific antibodies to multiple human antigens with minimal cross reactivity, due to their increased phylogenetic distance to humans, in comparison to mammalian hosts (Leonard *et al.*, 2007; Finlay *et al.,* 2011). By utilising whole PCa cells as the immunogen, the avian immune system was exposed to a wide range of antigens displayed in their native format. Two chickens were immunised, in case of a poor immune response, or other unforeseen experimental factors that may lead to the generation of a poor-quality antibody library. Following a full immunisation schedule, an immune response was observed in both chickens, shown by an antibody serum titre. The spleen, a rich source of antibody-producing cells (lymphocytes) was extracted from the both chickens, and the total RNA content was successfully isolated. The mRNA was then reverse transcribed to cDNA by first strand DNA synthesis, to generate a cDNA pool of each chicken's immune response. The cDNA pools were used as template DNA for the amplification of the variable genes, responsible for coding antibody specificity.

Chickens possess a simplified immune system in comparison to mammals, as their immunoglobulin repertoire derives from single V_H and V_L germ line sequences (Finlay *et al.*, 2006). The V_H and V_L genes from both chickens were amplified without the need for optimisation, using the method described by Andris-Widhopf *et al.* (2000). The variable genes were then spliced together with an SOE-PCR, yielding an scFv gene pool from both hosts. The libraries were transformed into the electrocompetent *E. coli* XL-1 Blue cells and their respective sizes were calculated. The ABG055 library (6×10^8 transformants) was over 300 times the size of the ABG048 library (1.8×10^6 transformants). While both libraries were deemed to be an acceptable size ($\geq 10^5$ transformants) to be used in antibody generation, larger sized libraries are more likely to contain desirable antibodies (Barbas III *et al.*, 2001). The ABG055 library was therefore selected to undergo enrichment against PCa cells through panning. Before proceeding with panning, the ABG055 library was checked for diversity. For a true estimation of library diversity, deep sequencing is required - in which 1000s of sequencing reactions must be carried out. This was not possible for practical reasons. Instead, small-scale sequencing was carried out to ensure the library was not made up of a high concentration of a single scFv gene. Comparisons of the CDR regions showed that none of the clones investigated were identical.

The ABG055 library was expressed as scFv-phage and underwent enrichment against PC3 cells via cell panning. A preliminary attempt at panning the library applied stringency by increasing the wash steps and reducing the number of cells between rounds. The number of washes increased from 3 to 8, while the number of cells decreased from 1×10^7 to 1×10^5, over 4 rounds. This approach yielded no significant round-to-round enrichment (data not shown) and was therefore modified. The second

attempt to pan the ABG055 scFv phage library kept the PCa cell number constant at 2 x 10^7 between rounds and introduced stringency by increasing the number of wash steps applied over the 4 rounds. A further modification involved plating the output phage-infected cells directly onto agar and allowing them to grow overnight. Following growth on plates, the cells were cultured in liquid media before helper phage infection. The purpose of including the agar-plating step was to reduce the risk of losing slow-propagating clones-of-interest. The result of this modified panning procedure was at least a tenfold enrichment of scFv-phage to PC3 cells. Colony-pick PCRs also showed that the scFv genes were present from the start to the end of the process.

The output scFv-phage from Rounds 3 and 4 were then selected for analysis in later chapters. Phage from both rounds were used infect non-suppressor *E. coli* Top10 F' cells to allow for soluble scFv expression, through recognition of the amber stop codon. This was desirable as bacteriophage are prone to non-specific binding of molecules (Popkov *et al.*, 2004; Jones *et al.*, 2016), and may cause interference in later screening experiments. More antibodies were selected from round 3 due to the higher polyclonal phage ELISA signal. Additionally, the risk of losing valuable binders amongst a background of rapidly growing clones increases with each round (Derda *et al.*, 2011). Of the 384 scFv selected, 96 were from the 4th round of panning, the remainder being from the 3rd round. Sequencing and CDR comparisons were carried out on 5 clones picked from round 4 output. Clones from round 4 were investigated over clones from round 3 as diversity is reduced with each round of panning (Derda *et al.*, 2011). While all 5 clones investigated had similar heavy chains, suggesting the panning did indeed apply a selective pressure on the scFv-phage, three unique scFv genes were present. The

resultant panel of scFv were carried forward for antigen-based characterisation in Chapter 4 and antigen-based screening in Chapter 5.

Chapter 4 – High density protein arrays for monitoring processes in recombinant antibody generation

4.1 Introduction

High-throughput technologies may help to further our understanding of, and capabilities in, immunisation-based recombinant antibody development. The aim of the work described in this chapter is to profile antibody-antigen interactions using protein array technology. The technology used employs large sequenced libraries of recombinant human proteins expressed in *E. coli* and chemically fixed to polyvinylidene difluoride (PVDF) membranes (Büssow *et al.*, 1998). Upon their inception, protein arrays were demonstrated to be useful for screening low numbers of recombinant antibodies from native libraries for rapid identification (Holt *et al.*, 2000). Since then, Kijanka *et al.* (2009a and b) have expanded on this by demonstrating the capacity to identify targets of both commercial monoclonal and polyclonal antibodies. Arrays may also be very valuable for defining the molecular targets of as yet uncharacterised antibodies and are also much better for characterisation of a range of antibodies currently poorly characterised in terms of their binding specificities. The technology has since been successfully applied to characterising the antibody profiles of sera isolated from patients with and without cancer, for identifying tumour-associated antigens (Kijanka and Murphy, 2009; Kijanka *et al.*, 2010; Brezina *et al.*, 2015). It has also resulted in the discovery of novel cancer biomarkers with both diagnostic and prognostic relevance (Fitzgerald *et al.*, 2013, 2015a; O'Reilly *et al.*, 2015). Implementation of protein array technology, may also allow the user to circumvent the traditional challenging antigen characterisation and identification processes that ultimately rely on peptide sequencing via MS.

In the research described here, protein arrays were employed to characterise scFv, the generation of which was described in chapter 3. Initially, the serum isolated from

chicken ABG055, before and after PCa cell immunisation, was analysed to reveal relevant derived immunogenic proteins. The scFv-phage library generated from the immune response was also characterised. Finally, the 384 scFv that were isolated following enrichment against PCa cells via panning, were pooled and added to protein arrays in a single experiment for rapid antigen characterisation. The library used was the Unipex library, consisting of 15,300 'in-frame' clones isolated from human foetal brain, T-cell and lung tissue cDNA, representing a total of 7,390 distinct human proteins (Source Bioscience, 2010). A cDNA library derived from a prostatic source may better suited to mimic the antigen pool contained within the PCa cell-derived immunogen, against which the avian scFv library was generated. However, the Unipex library has been successfully applied in the characterisation of the autoantibodies from patients with cancers originating other tissues than those the library was derived from, such as colon (Kijanka *et al.*, 2010). This is likely due to the large size of the Unipex library; displaying some level of antigen overlap with other tissue. Due to its size, the library is spotted over two PVDF membranes, with the proteins derived from the foetal brain tissue on membrane 1 and the remainder of the library on membrane 2.

The antigen lists generated from the protein array experiments were then examined, in collaboration with the Bioinformatics and Molecular Evolution Group of Dublin City University, to shed light on the antibody generation process. Chickens were employed as hosts in the research described in Chapter 3, due to their phylogenetic distance from humans in comparison to other traditional hosts, such as mice. This increased phylogenetic distance was expected to generate a maximal immune response, based on the difference in homology between the human-derived PCa-cell immunogens and the native avian proteins of the host. To elucidate the immunisation process, the Database

for Annotation, Visualization and Integrated Discovery (DAVID) tool (https://david-d.ncifcrf.gov/summary.jsp) was applied to analyse the antigen lists and determine if the immune response was directed towards antigens of specific cellular locations, processes, functions or pathways. The antigen datasets were also analysed for homology with their orthologs in chicken. Orthologs are genes in different species that have evolved from a common ancestral gene. It was expected that the immune response would be geared towards proteins with little-to-no similarity to their orthologs in chicken. Utilising these bioinformatic investigations, it was hoped that trends in the immunisation process, such as enrichment towards specific proteins, would be revealed. Knowledge of such trends would help to enhance our understanding of immunisation processes and would be useful for future antibody generation experiments.

4.2 Results

4.2.1 Western blot visualisation of an avian anti-PCa immune response

To visualise an avian-mounted immune response against human PCa cell proteins, serum taken from chicken ABG055 before and after a PCa cell immunisation regime was tested against PCa cell-protein-extract in a Western blotting format. PC-3 cells were lysed and proteins separated via SDS-PAGE, before transfer to a nitrocellulose membrane for Western blotting (see Section 2.4.1). Sera from ABG055 (both pre- and post- immunisation) were incubated with identical concentrations of cell lysate in western blot format (see Figure 4.1). Visualisation of binding events revealed a substantial increase in serum-antibody binding to PC-3 lysate proteins in the post-immunisation serum sample, as expected. A high background signal was also noticeable with TMB being developed uniformly across both membranes, suggesting both serum samples had heterophilic antibodies present that bound non-specifically to the membrane blocking reagent, milk marvel (MM) (Tate and Ward, 2004). Despite this background signal, differences in lysate protein-binding were still visible between the membranes.

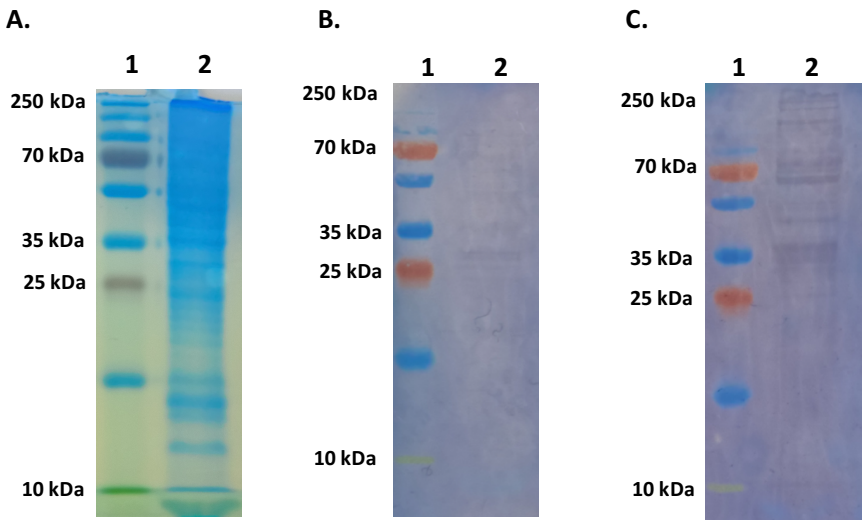

Figure 4.1 – Western blot analysis of an avian immune response to PCa cells. In all cases Lane 1 contains Page Ruler Plus Protein ladder for size reference and Lane 2 contains PC-3 cell lysate. Image A – SDS-PAGE of PC-3 cell lysate (50 μg) separated by SDS-PAGE. Image B – Western Blot analysis of avian serum IgY binding to PC-3 cell lysate before immunisation regime. Image C – Western Blot analysis of avian serum IgY binding to PC-3 cell lysate after immunisation. IgY-binding was detected using commercial rabbit anti-chicken IgY coupled with peroxidase-conjugated goat anti-rabbit IgG antibodies and the substrate was TMB.

4.2.2.1 Array profiling of secondary antibody false positives

The detection of the IgY, scFv-phage and scFv binding interactions on the protein array was based on a two- step sandwich procedure. The first antibody applied was to detect binding. The antibody used varied based on the source of the binding events to be visualised (IgY, scFv-phage or scFv). All primary antibodies were produced from rabbit hosts. This allowed the same detection antibody (AP-conjugated goat anti-rabbit IgG) to be utilised for visualisation of binding events. Due to the quantity of unique proteins on the arrays, some non-specific binding of secondary antibodies to the array was expected. To ensure these 'false-positives' were not mistaken to be binding events, each secondary antibody pair was profiled on the arrays in the absence of its manufacturer-listed target antigen (see Section 2.4.3). These false positives were excluded from the antigen-lists generated in this chapter. A total of 142 binding events were visualised across all arrays (see Tables 4.1, 4.2 and 4.3). Of these, 7 antigens were identified across multiple experiments, and hence were hypothesized to be caused by non-specific binding of the AP-conjugated goat anti-rabbit IgG antibody common to all immunoassays. The origin of the array test proteins appeared to affect the likelihood of non-specific binding. For example, the secondary antibodies selected for IgY profiling bound to a greater degree to array 1 (derived from foetal brain cDNA) than to array 2 (derived from T-cell and lung cDNA) (see Figure 4.2). The opposite was observed for the secondary antibodies selected to profile scFv (see Figure 4.3). The antibodies selected to profile scFv-phage showed the lowest level of cross-reactivity, as they only bound 2 antigens on array 1 and 5 on array 2. Removal of the false positives from the proceeding experiments was not considered to be detrimental, as they amounted to a small fraction of the total protein array population. The identities of false positives can be found in Section C of the Appendix.

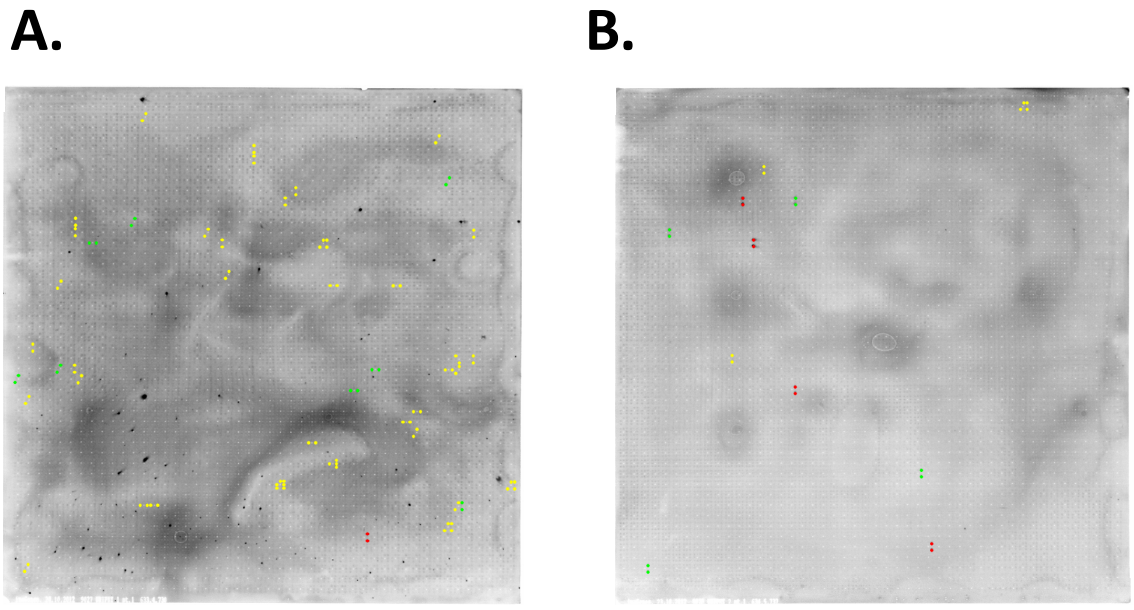

Figure 4.2 – Imagene software-identified binding events from array profiling of avian serum IgY secondary antibodies. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Binding signals are depicted as weak (yellow), moderate (green) and strong (red) intensities. Binding events were visualised using commercial rabbit anti-chicken IgY and AP-conjugated goat anti-rabbit IgG antibodies with AttoPhos AP Fluorescent Substrate. Raw image data are available in Appendix A.

Table 4.1 – Summation of the binding events in Figure 4.2.

Binding Events	Array 1	Array 2
Weak	42	4
Moderate	8	4
Strong	1	4
Total	51	12

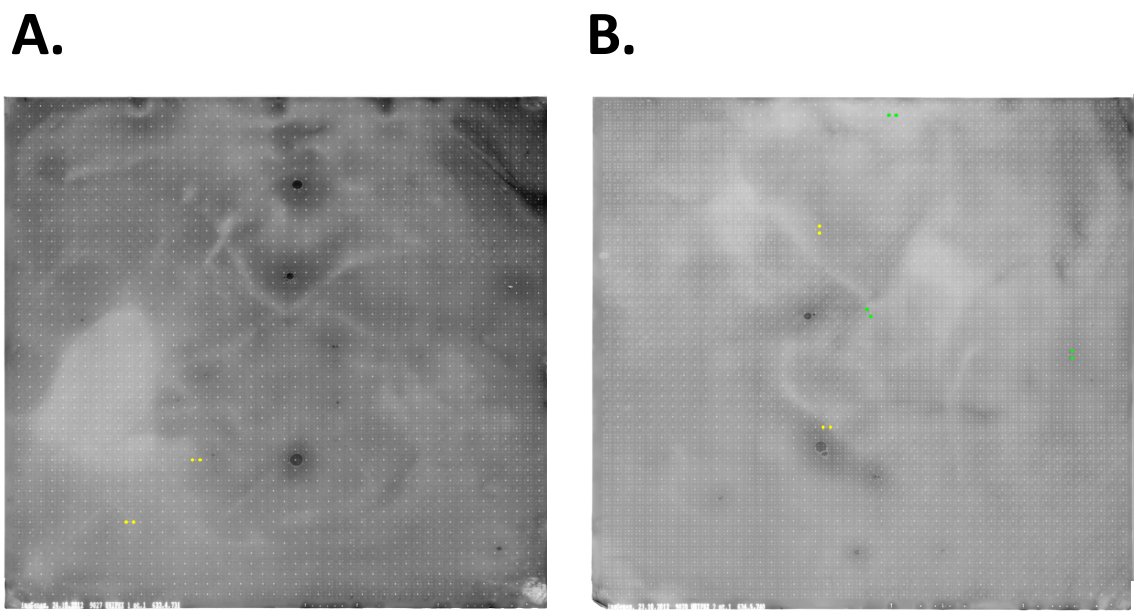

Figure 4.3 - Imagene software-identified binding events from array profiling of anti-bacteriophage secondary antibodies. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Binding signals are depicted as weak (yellow) and moderate (green). No strong (red) intensities were identified. Binding events were detected using a commercial rabbit anti-M13 + Fd bacteriophage coat proteins antibody coupled with an AP-conjugated goat anti-rabbit antibody and the fluorescent alkaline phosphatase substrate Attophos. Raw image data are given in Section A of the Appendix.

Table 4.2 – Summation of Binding events in Figure 4.3.

Binding Events	Array 1	Array 2
Weak	2	2
Moderate	0	3
Strong	0	0
Total	2	5

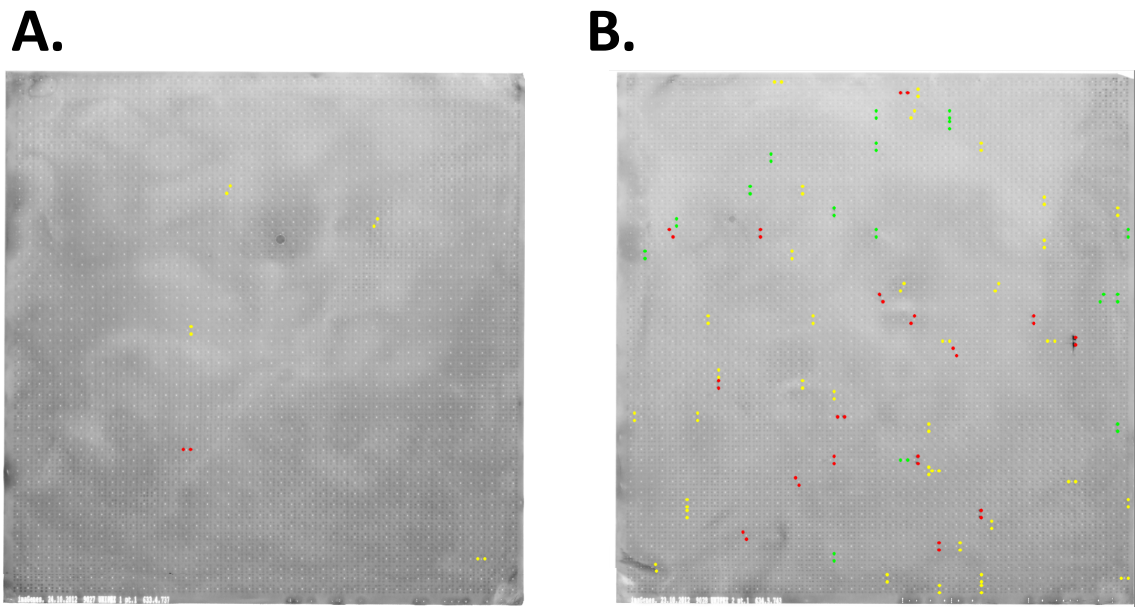

Figure 4.4 - Imagene software-processed image data obtained from control experiment profiling of scFv secondary antibodies. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Binding signals are depicted as weak (yellow), moderate (green) and strong (red) intensities. Binding events were detected using a commercial rabbit anti-HA antibody coupled with an AP-conjugated goat anti-rabbit antibody and the fluorescent alkaline phosphatase substrate Attophos. Raw image data are given in Section A of the Appendix.

Table 4.3 – Summation of binding events in Figure 4.4.

Binding Events	Array 1	Array 2
Weak	4	35
Moderate	0	16
Strong	1	16
Total	5	67

4.2.2.2 BLAST and SIM analysis of false positive subset

The false positive subset bound by the rabbit anti-chicken IgY and AP-conjugated goat anti-rabbit secondary antibodies were investigated for homology with chicken IgY and rabbit IgG. It was thought that non-specific binding may have been caused by homologous amino acid regions shared between the array-identified antigens and the immunogens to which the antibodies were raised. As the secondary antibodies were raised against chicken IgY and rabbit IgG, respectively, the false positive subset was compared with National Centre for Biotechnology Information (NCBI)-listed amino acid sequences for the constant and variable genes of both heavy and light chains, of the chicken IgY and rabbit IgG. For this, the Basic Local Alignment Search Tool (BLAST) was used to search for global amino acid out sequence overlaps between the identified antigens and each selected immunoglobulin gene. BLAST compares protein and nucleotide sequences for similarity, and assigns a value for the likelihood that a better match could be found by chance. The linear-space local similarity algorithm, SIM, was also used to search for short homologous regions between the false positives and the selected immunoglobulin sequences that may be missed by BLAST (see Methods Section 2.4.8). Eight of the 'false' positives identified were amino acid regions also present in immunoglobulins. All the antigens displayed some level of homology to at least one of the immunoglobulin amino acid sequences, reaffirming the profiling power of protein array technology. A summary of the BLAST and SIM analysis, presenting the highest scoring antigens, is presented in Table 4.4. The full table of antigen homology against each array-identified protein is available in Section B of the Appendix.

Table 4.4 – Summary of the 14 highest scoring protein array-identified antigens, when compared against four chicken and four rabbit immunoglobulin proteins for homology using BLAST and SIM alignment tools. Antigen comparisons that returned a BLAST E-Value or a SIM score equal to, or exceeding, set thresholds against one or more of the selected immunoglobulin genes were included. Threshold values for inclusion were a BLAST E-value of $\leq 1 \times 10^{-10}$ or a SIM score ≥50. Any value below the threshold is labelled as 'Below Threshold' (BT). A detailed table including all individual antigen-immunogen comparisons is available as Table 8.1 of Appendix B.

Clone Name	Chicken Immunoglobulin Genes				Rabbit Immunoglobulin Genes				Signal Intensity
	BLAST overlaps	BLAST E-value	SIM overlaps	Highest Score	BLAST overlaps	BLAST E-value	SIM overlaps	Highest Score	
PREDICTED: uncharacterized protein LOC749354 [Pan troglodytes]	1	4.00E-14	1	122	2	4.00E-148	2	1033	2
Clone SFV019_2F05H immunoglobulin heavy chain variable region	1	1.00E-48	2	393	2	3.00E-47	4	356	3
Clone IgA-MZ-aa42c-2 immunoglobulin alpha heavy chain variable region (IgA)	1	4.00E-30	2	210	1	1.00E-29	1	194	3
IGH mRNA for immunoglobulin heavy chain	2	2.00E-34	1	218	1	1.00E-41	2	291	3

VHDJ region, partial cds, clone:TRH1-16									
Immunoglobulin heavy constant gamma 1 (G1m marker)	1	3.00E-13	1	86	2	2.00E-125	2	890	3
Clone IP80 immunoglobulin heavy chain variable region	1	3.00E-44	1	313	1	3.00E-49	2	357	2
Immunoglobulin heavy constant gamma 1 (G1m marker)	1	6.00E-22	2	200	1	7.00E-23	2	213	2
Enhancer of polycomb homolog 1 (Drosophila)	0	BT	1	51	0	BT	0	BT	2
Inverted formin, FH2 and WH2 domain containing	0	BT	0	NA	0	BT	1	51	1
Single stranded DNA binding protein 4	0	BT	1	56	0	BT	0	BT	1
Rho GTPase activating protein 33	0	BT	1	54	0	BT	0	BT	1

Splicing factor 3b, subunit 4, 49kDa	0	BT	1	61	0	BT	0	BT	1
Protein phosphatase 1, regulatory subunit 26	0	BT	1	59	0	BT	0	BT	1
Alpha tubulin acetyltransferase 1	0	BT	1	50	0	BT	0	BT	1

Example: The AA sequence from highest ranked antigen 'PREDICTED: uncharacterized protein LOC749354' displayed very high AA sequence similarity to the immunogens. BLAST analysis of the antigen yielded an E-value (4×10^{-14}) below the set threshold (1×10^{-10}) against 1 of the 4 chicken immunoglobulin AA sequences, to which it was compared. SIM analysis comparing the same 4 immunoglobulin AA sequences to the clone, yielded similarity scores higher than the set thresholds (50) between the antigen clone to 2 of the immunoglobulin genes, with the highest score being 122. When the antigen AA sequence was compared against 4 rabbit immunoglobulin AA sequences, it achieved a BLAST similarity E-value below the set thresholds from 2 of the 4 comparisons, with the highest scoring overlap given an E-value of 1×10^{-148}. SIM analysis identified similarity between the antigen AA sequence and two of the rabbit immunoglobulin AA sequences above the set thresholds, with the highest similarity score being 1033.

4.2.3 Profiling an avian immune response to prostate cancer cells

The immune response of chicken ABG055 against PCa cell lines was characterised by protein arrays following the identification of detection-antibody-specific 'false-positives'. IgY-containing serum, taken before and after the immunisation regime, was profiled (see Methods Sections 2.4.4 and 2.4.7). A significant amount of background signal was observed as with Western blot analysis of serum IgY to PC-3 cell lysate. This was due to a number of factors. In addition to the heterophilic antibodies non-specifically binding to the milk protein (MM) blocking agent on the membrane, the chicken serum may have had IgY specific to bacterial proteins previously encountered by the chicken, which are likely to have increased the background by binding to the non-recombinant *E. coli* proteins present on the arrays. Analysis of both experiments identified 7 protein spots bound from the pre-immunisation bleed serum (see Figure 4.5 and Table 4.5) and 139 bound from the post-immunisation serum (see Figure 4.6 and Table 4.6). Only 1 false positive was identified in the post-immunisation sample, suggesting the secondary antibodies will bind their manufacturer-listed antigen in preference to the array-displayed proteins that share homology with the antigen. Of the 7 antigens identified in the pre-immunisation sample, 6 were also identified in the post immunisation profile. The antigen lists are available in Section C of the Appendix.

A. B.

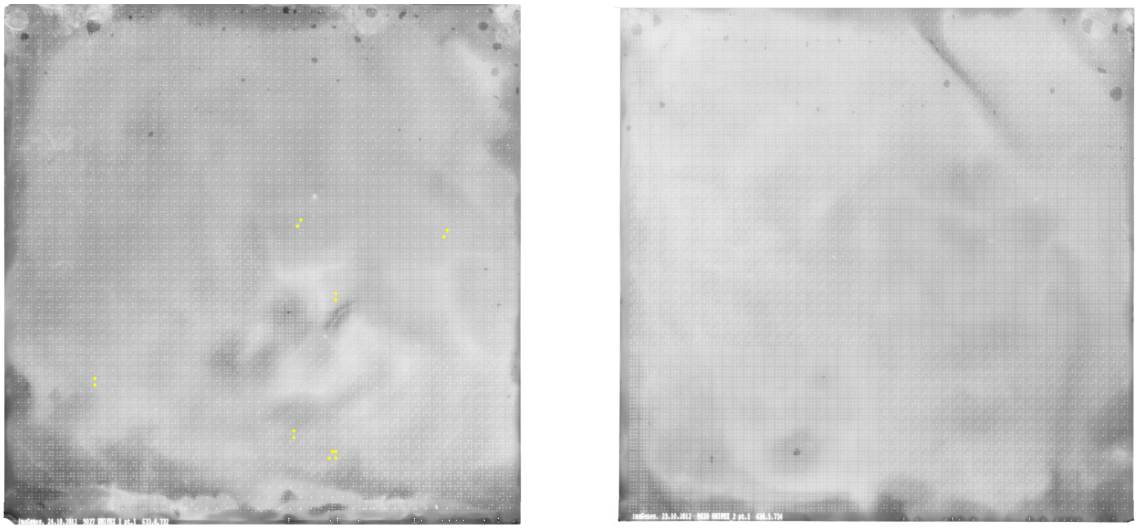

Figure 4.5 - Imagene software-processed image data obtained from array profiling of IgY in avian serum prior to a PCa cell immunisation regime. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Only weak signal intensities (yellow) were identified in this experiment. Binding events were visualised using commercial rabbit anti-chicken IgY and AP-conjugated goat anti-rabbit IgG antibodies with AttoPhos AP fluorescent Substrate. Raw image data is available in section A of the Appendix.

Table 4.5 – Summation of the binding events in Figure 4.5.

Binding Events	Array 1	Array 2
Weak	7	0
Moderate	0	0
Strong	0	0
Total	7	0

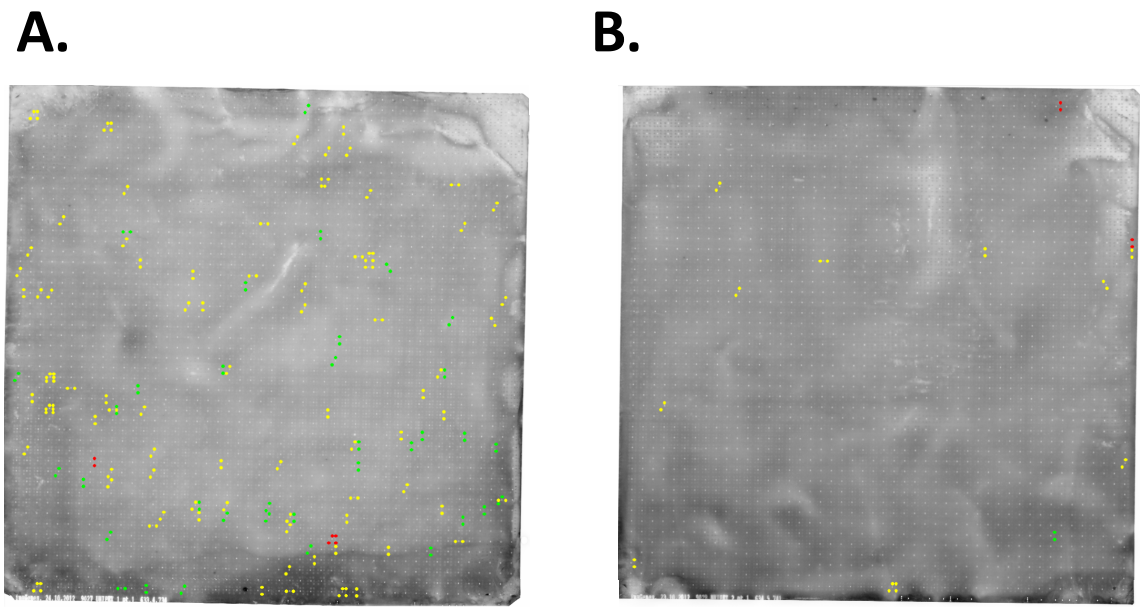

Figure 4.6 - Imagene software-processed image data obtained from array profiling of IgY in avian serum following a PCa cell immunisation regime. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Binding signals are depicted as weak (yellow), moderate (green) and strong (red) intensities. Binding events were visualised using commercial rabbit anti-chicken IgY and AP-conjugated goat anti-rabbit IgG antibodies with AttoPhos AP Fluorescent Substrate. Raw image data is available in Section A of the Appendix.

Table 4.6 – Summation of the binding events in Figure 4.6.

Binding Events	Array 1	Array 2
Weak	87	11
Moderate	35	1
Strong	3	2
Total	125	14

4.2.4 Antigen profiling of a scFv-phage library

Following characterisation of avian serum IgY profiles by protein arrays, the same approach was used to profile a scFv-phage library. The generation of the scFv-phage library was described in Chapter 3 using the cDNA isolated from chicken ABG055 following a PCa cell line immunisation regime. The ABG055 scFv-phage library was profiled (see Methods Sections 2.4.5 and 2.4.7) revealing a total of 16 antigens (see Figure 4.7 and Table 4.7), none of which were earlier identified as false positives. Two of the antigens identified were also present in the previously profiled post-immunisation immune repertoire. Like the serum IgY profiling, a high background signal was observed for the scFv-phage profiling. The scFv-phage antigen lists, as well as the antigens detected in multiple experiments, are available in Section C of the Appendix.

A. **B.**

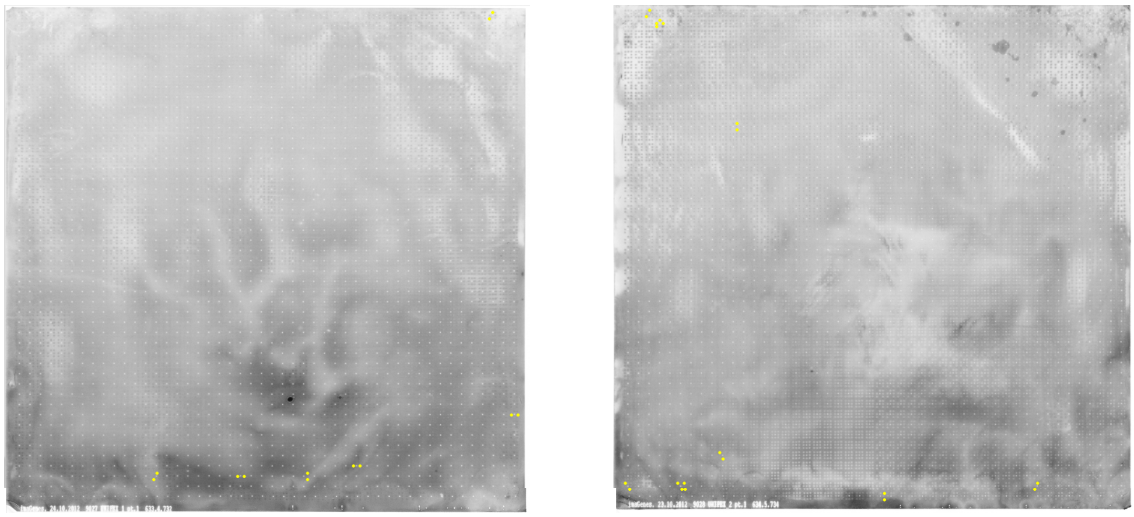

Figure 4.7 - Imagene software-processed images showing antigen profiling of an avian scFv-phage library generated against PCa cells using protein arrays. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Only weak signal intensities (yellow) were identified in this experiment. Following incubation of scFv-phage with array surfaces, binding events were detected using a commercial rabbit anti-M13 + Fd bacteriophage coat proteins antibody coupled with an AP-conjugated Goat anti-rabbit antibody and the fluorescent alkaline phosphatase substrate Attophos. Raw image are available in Sections A of the Appendix.

Table 4.7 – Summation of the binding events in Figure 4.7.

Binding Events	Array 1	Array 2
Weak	6	10
Moderate	0	0
Strong	0	0
Total	6	10

4.2.5 Antigen profiling of 384 scFv enriched against prostate cancer cells

Chapter 3 describes the generation of 384 avian scFv following enrichment experiments against PC-3 cells. The crude lysate of the scFv were pooled and analysed in a single experiment using Unipex protein arrays to reveal their antigen repertoire (see Sections 2.4.6 and 2.4.7). The results of this experiment showed 234 scFv binding events (see Figure 4.8 and Table 4.8), of which 5 overlapped with their false positive dataset and were subsequently excluded from the antigen list generated. A total of ten antigens bound by the scFv overlapped with previously generated profiles (9 were present in the post-immunisation antigen set and 1 was present in the scFv-phage antigen set). Unlike the IgY and scFv-phage profiling experiments, a low background signal was observed during the scFv profiling. A difference in trends was apparent between the anti-HA secondary antibody control experiment (see Figure 4.4) and the scFv experiment. Most binding events for the control profiling experiment occurred on Array 2. This was reversed when the scFv were included in the immunoassay. Lists of the identified and overlapping antigens are available in Section C of the Appendix.

A. B.

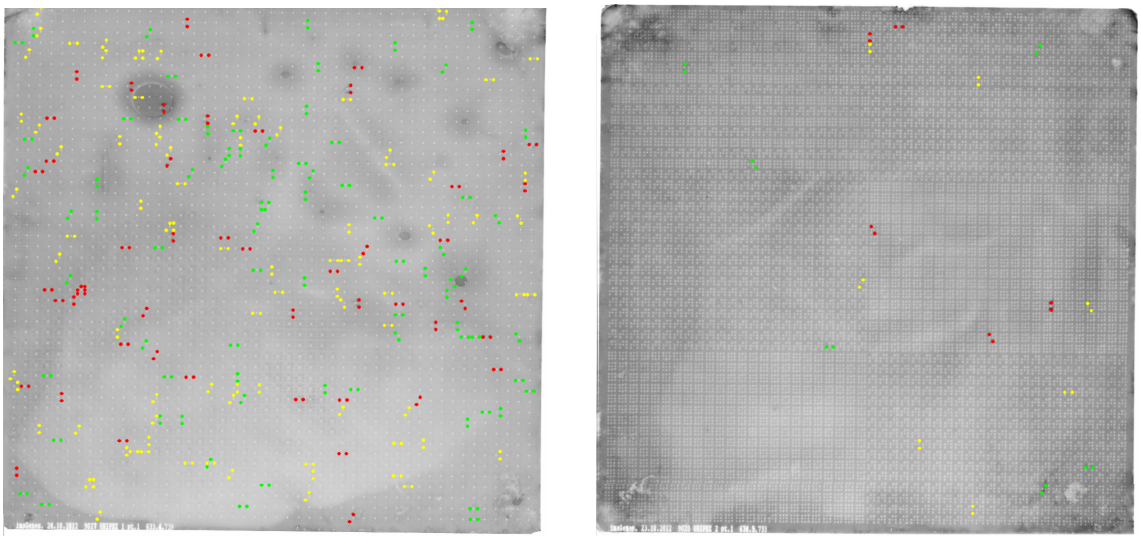

Figure 4.8 - Imagene software-processed images showing protein array profiling of scFv from 384 clones enriched against PCa cells via panning. Image A – Protein array 1 (derived from foetal brain cDNA). Image B – Protein array 2 (derived from lung and T-cell cDNA). Binding signals are depicted as weak (yellow), moderate (green) and strong (red) intensities. Following incubation of scFv-containing crude lysate with array surfaces, binding events were detected using a commercial rabbit anti-HA antibody coupled with an AP-conjugated goat anti-rabbit antibody and the fluorescent alkaline phosphatase substrate Attophos. Raw image data are available in section A of the Appendix.

Table 4.8 – Summation of binding events in Figure 4.8.

Binding Events	Array 1	Array 2
Weak	88	7
Moderate	76	6
Strong	50	5
Total	214	18

4.2.6 DAVID analysis of protein array-generated antigen profiles

The post-immunisation and scFv antigen datasets were converted into their Entrez Gene IDs (http://www.ncbi.nlm.nih.gov/gene) and analysed using the web-based Database for Annotation, Visualization and Integrated Discovery (DAVID) functional annotation tool (DAVID 6.8 Beta) (https://david-d.ncifcrf.gov/summary.jsp) (see Section 2.4.9). Here gene lists are loaded into the software to be sorted into function- based clusters and assigned enrichment scores. Cluster enrichment scores higher than 1.3 may be enriched , depending on the multiple statistical testing correction scoring (Bonferonni, Benjamini and False Discovery Rate) carried out by the software (Huang *et al.*, 2008). The aim of this experiment was to compare the antigen datasets generated to the complete proteome profile of the Unipex protein array library to determine if enrichment had occurred following the immunisation or panning procedures based on cellular location, process, function or pathway. The analysis of both scFv and post-immunisation antigen lists, however, revealed no statistically significant enrichment towards specific cellular location, process, function or pathway. A specific function-based enrichment was not expected from the avian post-immunisation antigen profile and is supported by this analysis. However, the antigen dataset of the scFv isolated following cell-panning was expected to be enriched towards cell surface antigens. The software identified a cluster of transmembrane proteins that appeared enriched, with an enrichment score of 4.5 (the highest score generated from this analysis), yet could not be considered statistically significant due to their associated multiple correction scores. Three additional clusters, one of which included regulators of transcription, scored above 1.3 from the scFv dataset (see table 8.12 in Appendix D). These were also ruled out as being statistically significant based on their multiple correction scores,

although may be worth investigating as potential cancer biomarkers in future studies. The output data from the DAVID analysis can be found in Section D of the Appendix.

4.2.7 Analysis of orthology of protein-array generated antigen profiles

A chicken was chosen as the host for immunisation with cell lines, partially due to its increased phylogenetic distance to humans, in comparison to other commonly used hosts. The expected outcome of the immunisation was a response to proteins that had little-to-no homology with the chicken's native proteins (Bowes *et al.*, 2011). To determine if this was the case, the post-immunisation IgY antigen profile of the chicken ABG055, was analysed for orthology with the chicken genome (see Section 2.4.9). The human genome, Unipex genome and scFv antigen profile were also included in the analysis for comparative purposes. The array-obtained antigen datasets from the pre-immunisation IgY and scFv-phage profiles were excluded due to their small size. The analysis indicated a trend in the post-immunisation response, towards orthologs that were conserved between both species. Only 10% of the post-immunisation highlighted antigens showed little-to-no homology with their orthologs in chickens, indicating enrichment of conserved orthologs (see Figure 4.9). In contrast, 73% of the human genome shared little-to-no homology with their orthologs in chicken, indicating that most orthologs shared between human and chicken are not conserved in nature. However, the Unipex gene set, like the post-immunisation dataset, showed a similar trend in enrichment towards highly conserved orthologs, with 24% of genes showing little-to-no homology. This was also the case with the scFv antigen set, with 15% of antigens showing little-to-no homology. The high density of conserved orthologs in the Unipex library was likely a contributing factor to the similar trend in conserved ortholog enrichment in the post-immunisation and scFv antigen profiles. However, the post-immunisation group still displayed the strongest trend towards conservation. It should be noted that due to the small post-immunisation and scFv sample sizes in comparison

to the human and Unipex genomes, the observed trend towards conserved orthologs cannot be considered statistically significant.

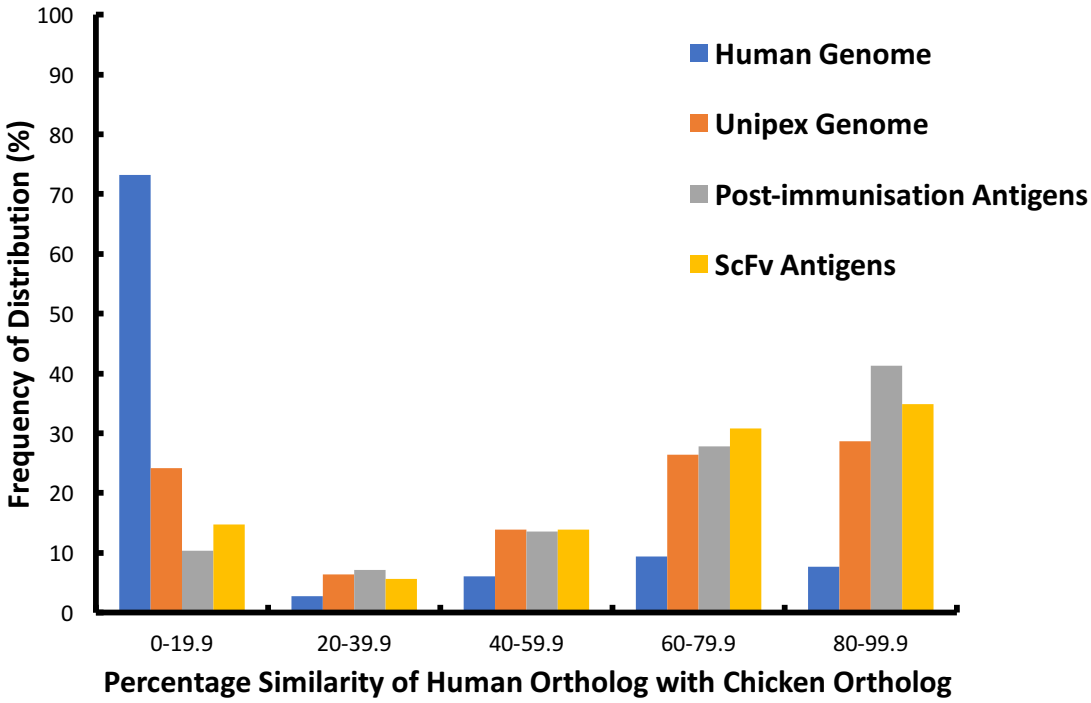

Figure 4.9 - Distribution of percentage similarity in human gene lists with their orthologs in chicken. The size of each data set varied - Human genome (65,576), Unipex genome (6,579), Post-immunisation antigens (126) and scFv antigens (143). The gene data sets were converted to a value of 100% and separated based on their percentage similarity with their chicken orthologs to provide a qualitative overview of trends.

4.3 Discussion

The aim of this chapter was to characterise the antigen profile at each stage of the recombinant antibody generation experiments, described in Chapter 3. To achieve this, Unipex protein arrays were employed to analyse the IgY serum profile of chicken ABG055, before and after a human PCa cell immunisation regime. The same analysis was performed on the scFv-phage library generated from the immune repertoire of chicken ABG055. Finally, the arrays were used to reveal the antigen profile of the 384 scFv isolated following panning-based enrichment of the scFv-phage library against PCa cells. The workflow of the process, with the observed number of binding events at each step, is illustrated in Figure 4.10.

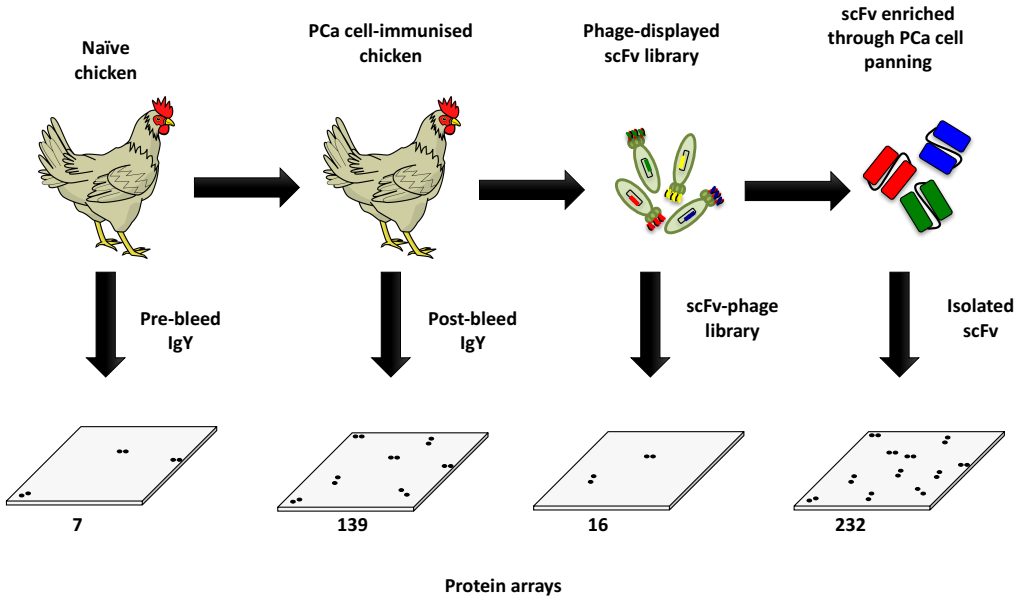

Figure 4.10 - Workflow overview results shown in Chapter 4. Protein array technology was employed to characterise the antigen profile of antibodies at sequential stages of the antibody generation process with the number of binding events visualised from each experiment. The samples analysed were: the serum IgY from a naïve chicken ABG055 (7 antigens bound), serum IgY from a chicken immunised against human PCa cells (139 antigens bound), the scFv-phage library generated following an avian immune response to PCa cells (16 antigens bound) and 384 scFv isolated from the scFv-phage library following PCa cell panning (232 antigens bound).

The need for experimental characterisation and validation of secondary antibodies prior to use is becoming increasingly emphasized in the immunoassay field (Uhlen *et al.*, 2016). Such validation experiments are flexible, but should preferably mirror the intended use of the antibodies. The secondary antibodies selected for use in the protein array experiments were polyclonal and, therefore, some level of non-specific cross-binding with the array might conceivably be expected. Therefore, before applying protein array technology to the antibody generation process, control experiments were carried out to profile each set of secondary antibodies, in the absence of their intended antigen, to determine their cross-reactivity.

The first experiment profiled the commercial rabbit anti-chicken IgY and AP-conjugated goat anti-rabbit IgG antibodies in the absence of chicken sera. This yielded a total of 63 bound antigens, which could subsequently be excluded from IgY-profiling experiments as false positives. Similar control profiling experiments were carried out before characterisation of the scFv-phage library and the scFv pool. Both assays utilised the same commercial AP-conjugated goat anti-rabbit IgG antibody selected for sera IgY profiling with a new primary antibody. For the scFv-phage control experiment, a commercial rabbit anti-Fd + bacteriophage coat proteins antibody was included, resulting in 7 binding events. The reduction in binding events, in comparison to the IgY profiling antibodies, may be due to the decreased homology between the manufacturer-stated immunogen, in this case (bacteriophage) and the human proteome. The scFv profiling control experiment included a commercial rabbit anti-HA antibody with the AP-conjugated goat anti-rabbit IgG antibody, yielding a total of 72 binding events. Interestingly, most of the positives were located on array 2. Array 2 of the Unipex array set was generated from cDNA isolated from T-cell and lung tissue, whereas array 1 utilised cDNA isolated from foetal brain tissue. All control experiment profiles were compared for overlap. A total of 7 antigens appeared in multiple datasets and were hypothesized to be non-specific binding from the AP-conjugated goat anti-rabbit antibody, as this antibody was utilised in each experiment. These false positive lists are available in Section C of the Appendix. The datasets produced from these experiments may be used as a false positive database for the commercial antibodies profiled and should be taken into consideration for future protein array or other immunoassay-based experiments. Such datasets may be improved upon in future experiments by utilising arrays that include additional portions of the human proteome and/or post-translational modifications. (Lemass *et al.*, 2016). For example, antibodies showing

higher rates of cross-reactivity may not be suited for use in complex immunoassays, such as immunohistochemistry, that are reliant on the specificity of the secondary antibodies for determining the presence of proteins-of-interest. Commercial monoclonal antibodies may offer improved specificity to the polyclonal antibodies used in these experiments. However, their inclusion was expected to increase the cost of each assay significantly compared to their polyclonal counterparts, due to the large volume of antibody reagent needed per array. In addition, additional protein arrays, the availability of which were limited, would be needed for validation of new secondary antibodies. Instead, the polyclonal antibodies already validated using scientifically acceptable methods (Uhlen *et al.*, 2016) were considered satisfactory and were shown to non-specifically bind to less than 2% of the Unipex array proteome.

The false positives identified by the antibody pair, selected for the chicken serum IgY profiling were investigated for sequence homology with their manufacturer-stated target antigens (IgY isolated from chicken and IgG isolated from rabbit). Here, the amino acid sequence of each array-identified antigen was obtained from NCBI databases, as well as amino acid sequences from different segments of chicken and rabbit immunoglobulins (variable and constant genes from both the heavy and light chains of each immunoglobulin). A comparative analysis of each false positive against the selected immunoglobulin sequences was carried out using BLAST and SIM alignment tools. Some level of similarity, with at least 1 of the selected immunoglobulin sequences, was observed for each array-identified antigen. Of the top ten-ranked antigens (based on similarity scoring), 8 were identified as sequences derived from immunoglobulin proteins. The level of homology between these antigens, with the immunoglobulin gene

sequences selected, illustrates ability of protein array technology to accurately characterise antibodies.

Following the profiling of secondary antibodies and subsequent exclusion of the false positives from the experimental lists, a total of 417 recombinant proteins were identified to be bound by IgY, scFv-phage or scFv over all experiments. These included 117 duplicate binding events, in which multiple clones present on the array expressed different segments of the same protein and were bound simultaneously (available in Section C of the Appendix). Removal of the duplicates yielded 300 unique antigens, and of these, 18 were present in more than one dataset (see Section C of the Appendix). A table summarising this non-redundant set of antigens across all four experiments is shown below (Table 4.9). Furthermore, a Venn diagram summarising the overlap between each antigen dataset is shown in Figure 4.11.

Table 4.9 – Summation of non-redundant antigens bound by antibodies across all four profiling experiments.

Profiling experiment	Redundant clones	Excluded 'false-positives'	Non-redundant antigens
Pre-immunisation	7	0	7
Post-immunisation	139	1	129
scFv-Phage	16	0	16
scFv	232	5	166

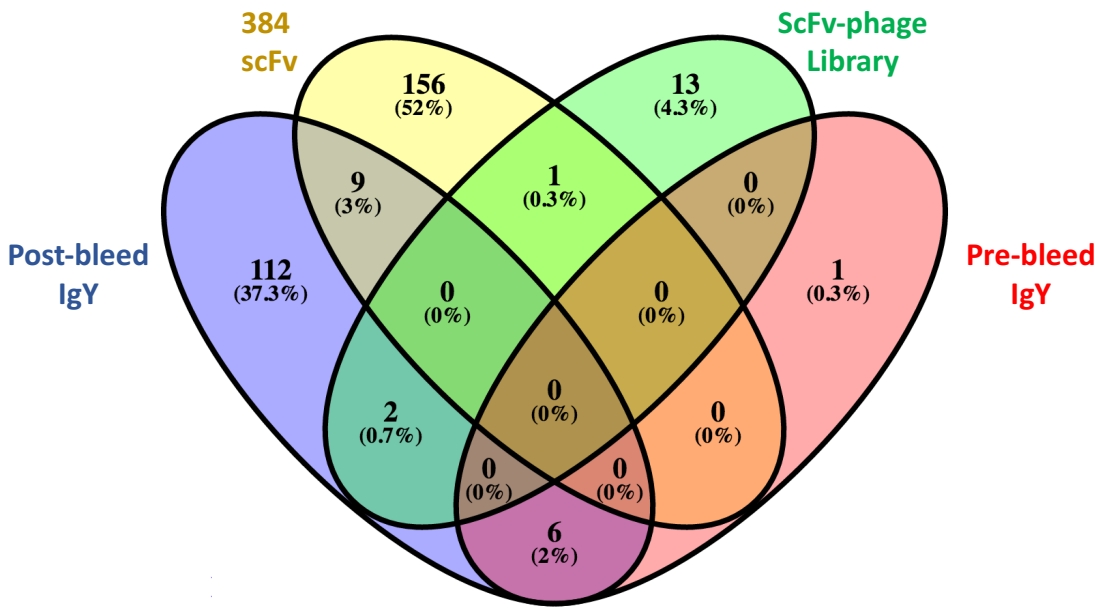

Figure 4.11 – Venn diagram showing the overlap between antigens found across all four datasets.

The array profiling studies on the serum isolated from chicken ABG055 were consistent with the Western blotting experiments, in that an increase in binding to human protein was apparent following the immunisation regime. In addition, a high level of background

signal was observed in both protein array and Western blot experiments, suggesting the serum also contained immunoglobulins that bound the blocking agent MM. Despite this background, 7 unique binding signals were identified from the pre-immunisation serum and 129 were identified from the post-immunisation serum. Of the 7 antigens identified in the serum before the immunisation regime, 6 were also present in the post-immunisation serum. The Unipex array consists of human foetal brain, T-cell and lung cDNA libraries spotted onto two PVDF membranes. Array 1 exclusively houses cDNA isolated from the foetal brain cDNA. A trend was observed that the majority of binding events in all experiments occurred on Array 1. The antigen overlap from each dataset was influenced by the stage of recombinant antibody generation. For example, 6 of 7 antigens present in the pre-immunisation profile overlapped with the post-immunisation profile, but not with the scFv produced at the end of the process. This can be explained by the chicken retaining its immunity to the original antigens over the course of the immunisation regime. However, as these antigens were bound before the process they are unlikely to have any link to the targeted immunogen (PCa cells). Hence, immunisation and panning of the scFv-phage library was expected to enrich for antigens present in PCa cells.

A major observation of this research was that the antigen profile of the antibodies investigated (IgY, scFv-phage and scFv) changed as the antibody generation process progressed. The majority of antigens detected over the course of these experiments were not identified in more than one dataset, despite the scFv being derived from the immune repertoire of the chicken following immunisation. This may be explained by the scFv, having undergone experimental enrichment against PCa cells via panning in addition to being isolated and expressed in monoclonal, were likely present at higher

concentrations than the parent antibodies they were derived from (in the post-immunisation and scFv phage datasets). Despite mounting an immune response against PCa cells, the IgY of the chicken serum was polyclonal and expected to be highly diverse in terms of antigen specificity (Bowes *et al.*, 2011). In addition, the high background signal observed during the IgY and scFv-phage profiling experiments may have masked weaker binding signals of antibodies present in low concentrations. Repetition and optimisation of experimental conditions may yield improvements to array image output as well as the reliability of the output antigen lists. However, this was not possible here, due to the limited availability of protein arrays and necessary regents such as sera. An overall assessment of the profiling experiments carried out in this chapter would rank scFv-phage as the least suitable antibody format to profile on arrays, and scFv as the best.

The profiling of an immune response may direct the researcher to antigens that are more likely to produce specific antibodies through application of immobilised panning. Selecting an antigen observed in an immune response and using that same antigen to pan the resultant scFv-phage library would be more likely to produce a specific antibody and may be another viable application of protein arrays. In addition, knowledge of how a host responds to complex protein mixtures may have implications for selecting specific antigens and hosts at the start of future antibody generation processes (Finlay *et al.*, 2005). The observation of trends, such as enrichment towards proteins with a specific function or location, or proteins that have reduced homology with the native proteome of the host, would be useful in evaluating the candidate antigens suitability for producing an antibody. Here, the post-immunisation and scFv antigen datasets were analysed for function-based and orthology-based enrichment to further elucidate the

antibody generation process. DAVID analysis was employed to compare the data sets generated in the development process with the Unipex library. No enrichment towards cellular location, process, function or pathway of statistical significance was detected in the post-immunisation or the scFv data sets. This data suggests that the chicken ABG055 mounted a general, non-specific immune response to the proteins present in the cells. There was an expectation, however, that the cell panning process would enrich towards plasma membrane proteins and although membrane proteins were identified as a potentially enriched cluster, it could not be confirmed from the multiple correction scores. While antibodies have been produced to internal proteins using a similar methodology to this research (Shukla and Krag, 2010; Zhou *et al.*, 2012), the majority of studies isolate antibodies to membrane proteins (Popkov *et al.*, 2004; Even-Desrumeaux *et al.*, 2013; Keller *et al.*, 2015). It may be that internalisation of scFv-phage is more common than thought during cell panning, but is not reported as the majority of screening approaches focus for cell membrane-binding candidates. There are a number of possible routes to non-receptor mediated internalisation, such as micropinocytosis, fast endophilin-mediated endocytosis, clathrin-dependent and independent endocytosis and fluid-phase uptake (Pavelka and Roth, 2010; Lim and Gleeson, 2011; Mayor *et al.*, 2014). It is also possible that scFv specific to internal antigens were carried through the rounds of the panning process by phage binding non-specifically to the cytoplasmic membrane of the cells (Jones *et al.*, 2016).

Despite not showing statistically significant enrichment of specific functions, locations, processes or pathways, the DAVID tool provided a function-based perspective on each antigen dataset that may otherwise be missed by individual interpretation or bias. For example, the top pathway cluster the software recognised from the post-immunisation

dataset, with an enrichment score of 1.0, was 'regulation of the actin cytoskeleton', with 6 genes implicated in the pathway present in the antigen dataset. Changes to this regulatory process have been linked to cancer cell migration and invasion (Yamaguchi and Condeelis, 2007) and isolation of the antibodies that bind these 6 proteins may be achievable with a traditional panning and screening approach. Antigens implicated in regulation of transcription were identified as one of the top trends in enrichment in the scFv dataset, with an enrichment score of 1.5. Alterations to transcription regulation are logically relevant to cancer pathogenesis and antigens present in this dataset have been highlighted as such. For example, expression of the antigen Jumonji Domain Containing 1C has been shown to play a critical role in acute myeloid leukemia survival (Chen et al., 2015). Hence, the antigens present in clusters that have a phenotype with a relevance to cancer, may be investigated for their role in PCa in future studies.

The same antigen lists were analysed for trends in orthology-based enrichment. The phylogenetic distance between human and chicken is larger than other standard hosts such as mouse and rabbit. Selecting chicken as a host, to mount a response against a complex mixture of human-derived proteins, was expected to increase the immunogenicity of the human proteins due to this distance. There was an expectation of finding enrichment to human proteins that were not homologous to those in chicken (Bowes *et al.*, 2011). To determine if this was the case, the antigen datasets from post-immunisation and panning profiles were again compared to the Unipex library, this time using the data-mining Biomart tool. The human genome was also included for comparative purposes. Each antigen was linked to its Ensembl gene ID in human and subsequently compared to its ortholog, or orthologs, in chicken. A strong difference in trends of homology between orthologs was first observed between the Unipex library

and the human genome. In the human genome >70% of genes were found to share little-to-no homology with their orthologs in chicken. This trend was reversed in the Unipex library, with >70% of all genes showing >20% homology with their orthologs. This trend of enrichment towards homologous orthologs was further increased in the post-immunisation dataset, conflicting with the current consensus that the avian immune response is primarily directed towards mammalian proteins that share minimal homology to native chicken proteins. However, the contrast in the dataset sizes used to carry out these comparisons meant that statistically significant conclusions could not be drawn, yet the immune response did qualitatively appear to be enriched towards highly conserved orthologs. This finding would be strengthened if observed in a follow up study utilising a protein array that better matched the human genome in terms of ortholog conservation.

In summary, this chapter utilised protein array technology to highlight and analyse the antigen profile at each stage of the recombinant antibody generation process. Characterising the antigen profile, particularly of a host-mounted anti-cancer immune response, or antibodies enriched against a cancer cell line via panning, was considered to be the most advantageous application of the technology. The following chapter employs traditional immunoassay screening techniques to determine if the scFv specific to their array-identified antigens could be isolated.

Chapter 5 – Characterisation of scFv clones for protein array-identified antigens

5.1 Introduction

The use of cell panning in phage display technology is an excellent method to generate diverse pools of recombinant antibodies against both known and unknown cancer-associated antigens. The major drawback, however, is the inability to distinguish antibodies to antigens-of-interest, from ubiquitous antigens in the earlier stages of the process. As a result, antibody pools generated with cell panning technology require extensive characterisation to determine the identity of the corresponding antigens (Siva *et al.*, 2008; Keller *et al.*, 2015). Without prior knowledge of the identity of the antigen, this labour-intensive approach presents the investigator with a high risk-to-reward work-flow. The isolation of antibodies to previously well-characterised antigens, to which there are already established and commercially-available antibodies, would negate the overall value of such experiments. Hence, isolated antigen immunisation coupled with fixed-antigen screening is still used as a more reliable method for the generation of valuable antibodies (Leonard *et al.*, 2007;Conroy *et al.*, 2012; Shih *et al.*, 2012).

The goal of the work described in this chapter was to validate the protein array-based scFv antigen identification experiment described in Chapter 4. This was to be achieved through the isolation of the scFv corresponding to a subset of array-identified antigens. Here, four antigens were selected for screening experiments based on being either uncharacterised or confirmed as relevant in the PCa disease state. SET Domain Containing 2 (SETD2) (also known as hypoxia-inducible factor 1) is a transcription factor that controls the expression of multiple proteins implicated in angiogenesis, erythropoiesis, glycolysis, and invasion pathways in cancer cells. Androgens have been shown to increase SETD2 expression in androgen-dependent PCa cell lines and

overexpression of the protein has been associated with increased cancer aggressiveness – highlighting SETD2 as a potential biomarker and therapeutic target in PCa (Mabjeesh *et al.*, 2003). Proteasome Subunit Alpha Type-7 (PSMA7) is a multicatalytic proteinase complex involved in cell cycle regulation and has been shown to enhance androgen receptor activity in prostate cancer cells (Lin *et al.*, 2002). Its value as a potential therapeutic target or biomarker of metastasis has been highlighted (Hu *et al.*, 2008). The third selected antigen, Zinc Finger Protein 358 (ZNF358), is relatively uncharacterised in PCa and other cancers to date. The original publication describing the protein hypothesized it to be a nuclear transmembrane hormone receptor, yet the emphasized the need of further investigation of the protein (Wride *et al.*, 2002). There are currently only polyclonal antibodies available for ZNF358 (Antibodypedia, 2016). Therefore, the production of a recombinant antibody would be useful for investigating the relevance of the protein in PCa. The final antigen selected for screening was the ligand-dependent transcriptional regulator Retinoic Acid Receptor Gamma (RARG). Aberrant expression of RARG has been linked to multiple types of cancer and has been shown to be a valuable immunohistochemistry biomarker in differentiating prostate cancer from benign prostatic hyperplasia and prostatic intraepithelial neoplasia based on its cellular localisation (Richter *et al.*, 2002).

In this chapter the SETD2, PSMA7, ZNF358 and RARG-expressing *E. coli* clones from the Unipex library were acquired and their encoded antigens were expressed, purified and sequenced for screening experiments. The strategy employed here was to screen the 384 scFv produced in Chapter 3 against each antigen via monoclonal ELISAs. ScFv observed to specifically bind the antigens were then to be tested in Western blot format. Following screening, further analysis was performed on one clone, labelled P3P2F10,

which showed selective binding to ZNF358 during the initial screening experiments and warranted further characterisation. This scFv was expressed and purified by IMAC for downstream investigations. The purified scFv was observed to bind ZNF358 both in ELISA and Western blot formats. This binding interaction was then kinetically characterised using a Biacore 3000 in a label-free format. Overall, this chapter showed that protein array-characterised scFv can be mapped to their corresponding antigens using a traditional immunoassay-based screening approach.

5.2 Results

5.2.1 Primary screening experiments

5.2.1.1 Sequencing of antigens and predicted protein size

The acquired clones, encoding PSMA7, SETD2, ZNF358 and RARG were sequenced to determine whether they encoded complete or partial fragments of their manufacturer-listed antigens. The obtained amino acid sequences were matched to their NCBI-listed sequences using BLAST and aligned with the Clustal Omega alignment tool (see Section 2.5.1.2). Following alignments, the expected size of each encoded antigen fragment was estimated using the ExPASy molecular weight calculator tool (see Table 5.1). In all cases, the proteins encoded by the clones matched their Unipex database-listed identities (see Figures 5.1, 5.2, 5.3 and 5.4). The presence of a vector-encoded His_6 region confirmed each antigen was suitable for purification by IMAC.

Figure 5.1 – Sequence alignment of the clone IMGSp9027C0934D (labelled UNIPEX) with SETD2 (NCBI Reference Sequence: NP_054878.5) (labelled SETD2). Note: SETD2 is a 2,564 AA protein, here only 344 C-terminal AA are displayed. The Unipex clone matched the C-terminal end of the protein, rather than the full-length protein. The presence of the His_6 tag indicated that the clone was suitable to be affinity-purified.

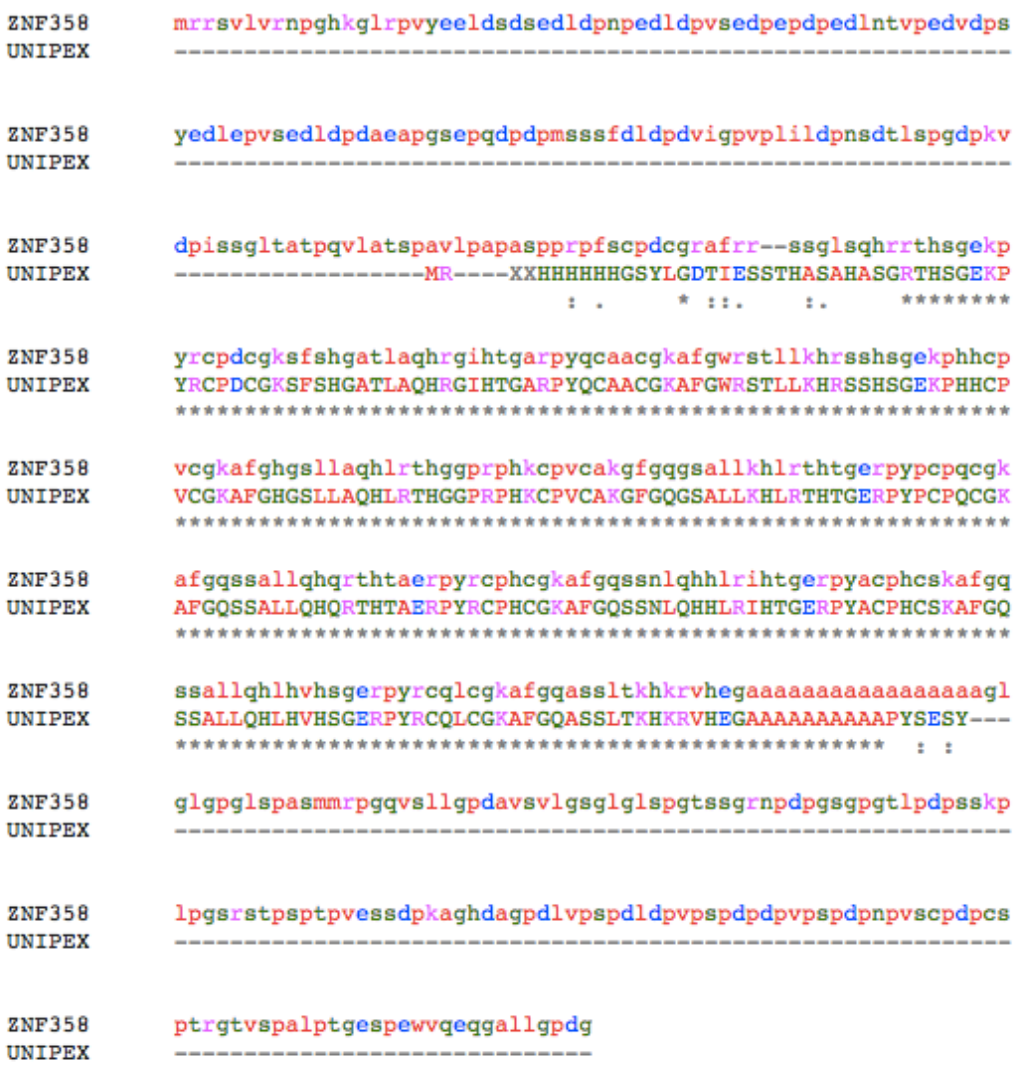

Figure 5.2 – Sequence alignment of the clone IMGSp9027B1237D (labelled UNIPEX) with ZNF358 (NCBI Reference Sequence: NP_060553.4). The full-length wildtype protein is 568 AA. The Imagene clone codes a fragment of the total protein – overlapping with 361 AA in the centre of the protein. A His_6 tag is also present at the start of the sequence indicating it is suitable for IMAC purification.

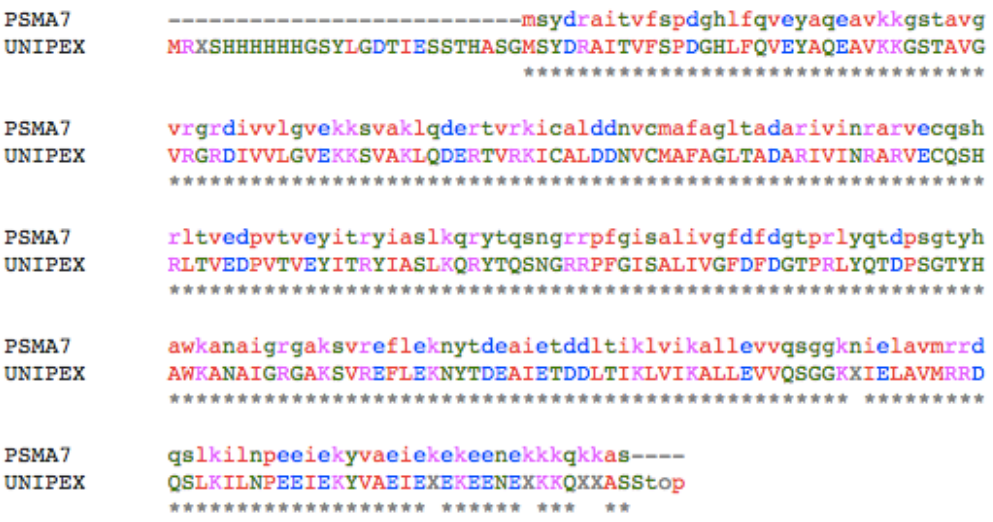

Figure 5.3 – Sequence alignment of the clone IMGSp9027C1115D (labelled as UNIPEX) with PSMA7 (NCBI Reference Sequence: NP_002783.1). The clone encodes the complete 248 AA wildtype protein as well as a His_6 tag making it suitable for purification.

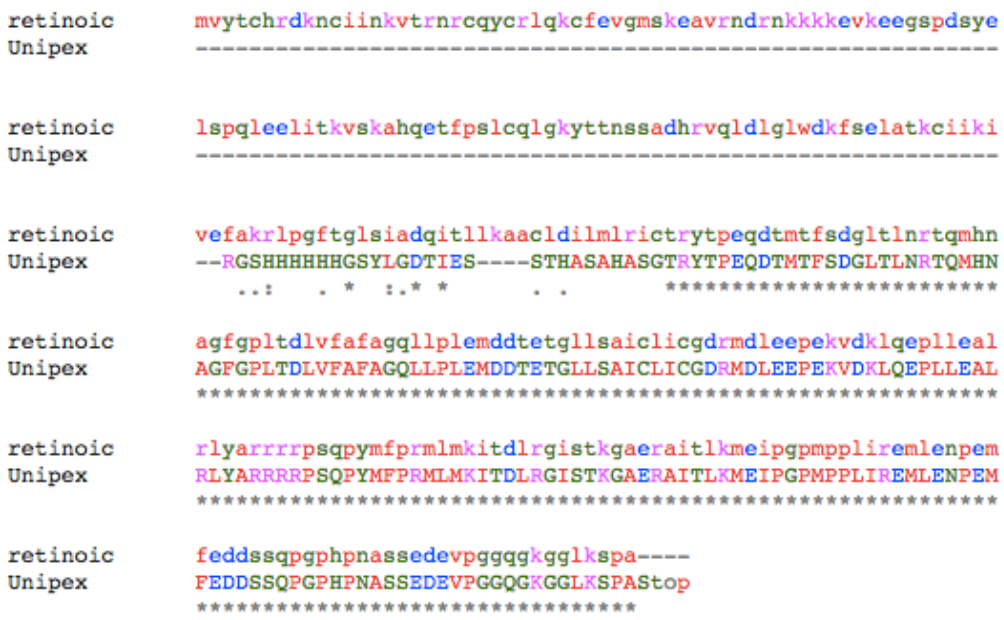

Figure 5.4 – Sequence alignment of the clone IMGSp9027C0873D (labelled as Unipex) with RARG (NCBI Reference Sequence: NP_001230660.1) (labelled as Retinoic). The Unipex clone encodes a fragment overlapping with the C-terminal of the 333 AA protein in addition to a His_6 tag, making it suitable for screening experiments.

Table 5.1 – ExPASy Molecular weight calculator-estimated sizes of protein fragments.

Protein Origin	Imagene ID	Estimated Size (Da)
SETD2	IMGSp9027C0934D	31,829
ZNF358	IMGSp9027B1237D	29,730
PSMA7	IMGSp9027C1115D	30,919
RARG	IMGSp9027C0873D	23,376

5.2.1.2 Expression and purification of protein array-identified antigens for screening

Following confirmation of antigen identity, the proteins were expressed in 100 mL cultures and purified by IMAC to produce stocks sufficient for downstream screening and characterisation (see Section 2.5.1.1). The IMAC elution and wash fractions were analysed by SDS-PAGE (see Section 2.3.23) for the presence of bands, equivalent to the previously estimated fragment sizes, that were also bound by an anti-His_6 detection antibody in a Western blot (see section 2.3.24). The ExPASy Molecular weight tool provides a theoretical size only and, therefore, inaccuracies between the expected size and observed size <20 kDa could be expected. All clones (SETD2, PSMA7, ZNF258 and RARG) yielded protein bands of a similar size to the ExPASy estimates (see Figures 5.5 and 5.6). In all cases, the single dominant protein bands were also bound by a peroxidase-conjugated anti-His antibody to the relevant antigen in Western blotting format and thus they were considered suitable for use in screening purposes.

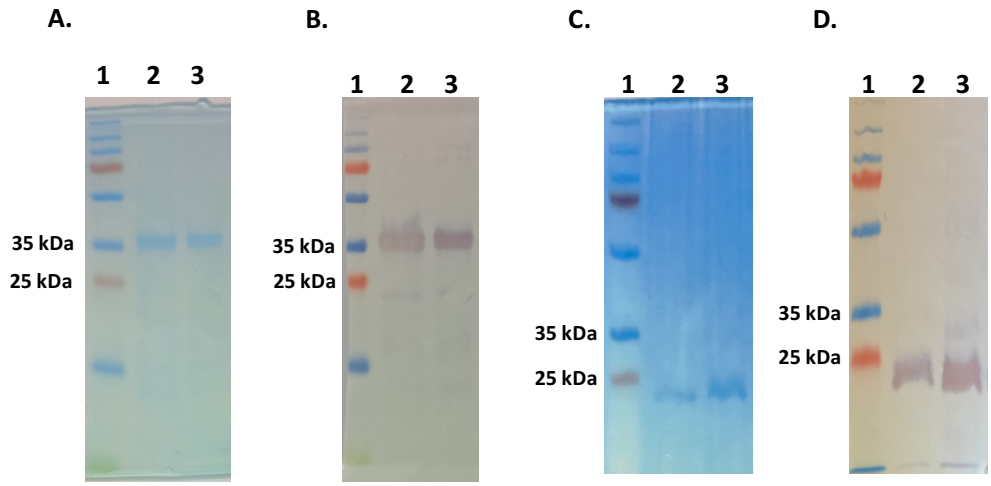

Figure 5.5 - SDS PAGE gels (A and C) and Western blots (B and D) of purified SETD2 (A and B) and RARG antigens (C and D). In all images, lane 1 contains a protein ladder for size determination, lane 2 contains a sample from the IMAC wash fraction and lane 3 contains a sample of the eluted fraction. Antigen purification was confirmed in all cases by the presence of a dominant band in close proximity to each of the antigen's estimated fragment size, that was also bound by a commercial HRP-conjugated anti-His_6 antibody and visualised using the substrate TMB. The estimated size of SETD2 was ~32 kDa, the fragment migrates just above the 35 kDa marker. RARG migrates to a positon just below 25 kDa which is in line with its estimated size of ~23 kDa.

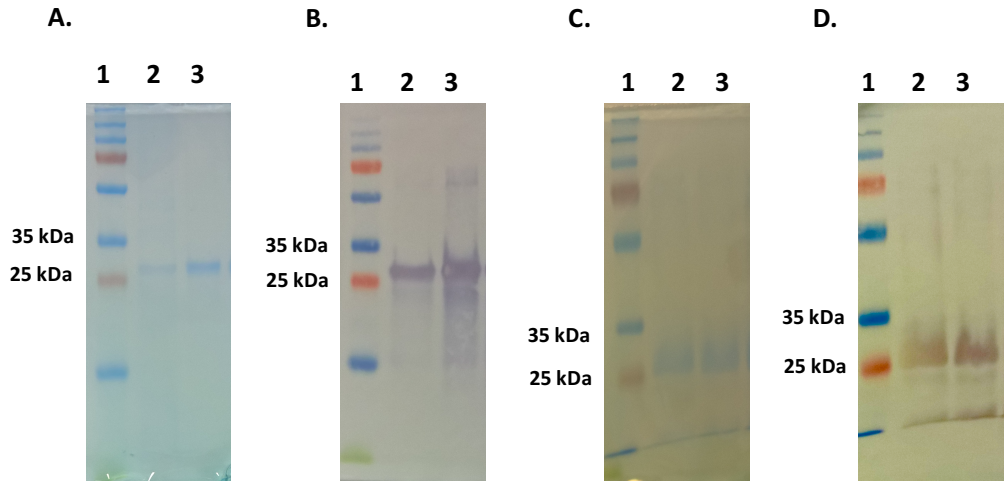

Figure 5.6 - SDS PAGE gels (A and C) and Western blots (B and D) of purified ZNF358 (A and B) and PSMA7 (C and D) antigens. In all images, lane 1 contains a protein ladder for size determination, lane 2 contains a sample of the wash fraction and lane 3 contains a sample of the eluted fraction. Antigen purification was confirmed in all cases by the presence of a dominant band in the SDS gel, in proximity to the estimated fragment size, that was also bound by a commercial HRP-conjugated anti-His_6 antibody and visualised using the substrate TMB in a Western blot format. The estimated size of the ZNF358 fragment was ~30 kDa, which is consistent with images A and B. The PSMA7 fragment was estimated to have a size of 31 kDa, which is consistent with the images C and D.

5.2.1.3 Screening of 384 scFv against selected antigens by monoclonal ELISA

Following purification and validation, the antigens were used to screen the 384 scFv described in Chapter 3. Due to the presence of urea in the IMAC elution buffer, which, at high concentrations, can interfere with protein quantification, antigen concentrations were determined at a 1/10 dilution using the BCA assay (see protocol in Section 2.5.1.3). Initial scFv screening was carried out by indirect ELISA, in which the scFv-containing lysates were tested for binding against antigen-coated wells (see Figures 5.7, 5.8, 5.9 and 5.10). Clones that yielded an absorbance >3 times the value of the highest negative control were considered as potential binders. The controls investigated for non-specific binding by excluding, respectively, the antigen, scFv and secondary antibody from the immunoassay. During PSMA7 screening there was significant non-specific binding of the antigen by the commercial HRP-conjugated anti-HA antibody, resulting in a cut-off higher than any scFv binding signal isolated from panning round 3 (see Figure 5.9). PSMA7 was therefore, retrospectively considered to be a false positive and excluded from downstream screening experiments. Multiple clones yielded binding signals to SETD2 (Figure 5.8), RARG (Figure 5.7) and ZNF358 above the applied cut-off values (Figure 5.10). To identify specific scFv, the binding signal of the clones to the antigens SETD2, RARG and ZNF358 were compared (see Figure 5.11). This allowed for discrimination between antibodies displaying specificity (for example the clone P3P2F10) and heterophilic scFv. It was observed that a significant number of scFv from panning round 4 were non-specific, yielding high signals against all antigens, for example P4P1D4.

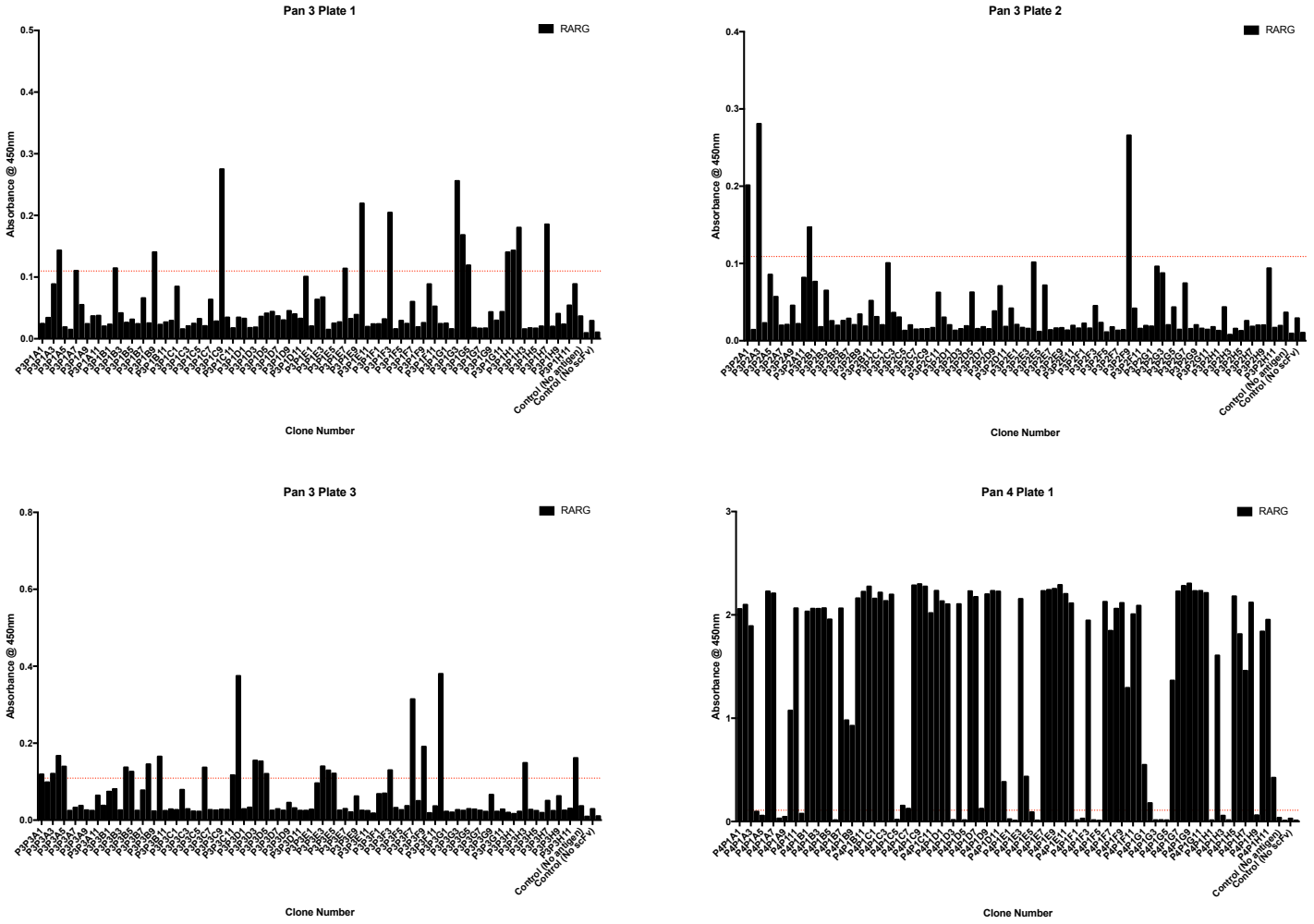

Figure 5.7 – Soluble monoclonal indirect ELISA screening for anti-RARG scFv. Clones (384 in total) were selected from PCa cell panning rounds 3 and 4. The scFv-containing supernatant was incubated on RARG-coated plates and binding was detected using a commercial HRP-conjugated anti-HA antibody and the substrate TMB. Each graph included a background cut-off that was 3 times the highest negative control value (indicated as a red horizontal line) to highlight scFv binding.

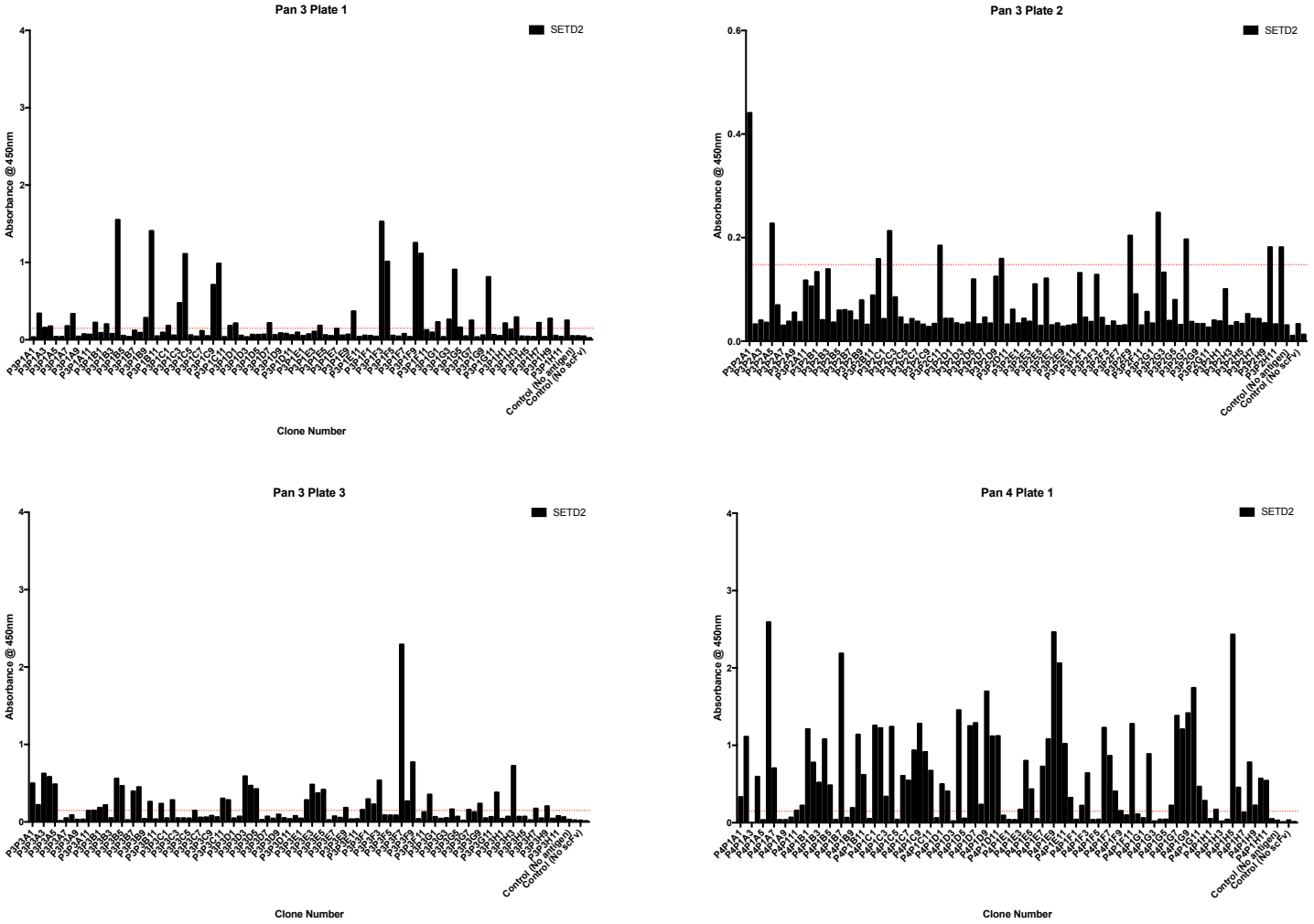

Figure 5.8 – Soluble monoclonal indirect ELISA screening for anti-SETD2 scFv. Clones (384 in total) were selected from PCa cell panning rounds 3 and 4. The scFv-containing supernatant was incubated on SETD2-coated plates and binding was detected using HRP-conjugated anti-HA antibody and the substrate TMB. Each graph included a background cut-off that was 3 times the highest negative control value (indicated as a red horizontal line) to highlight scFv binding.

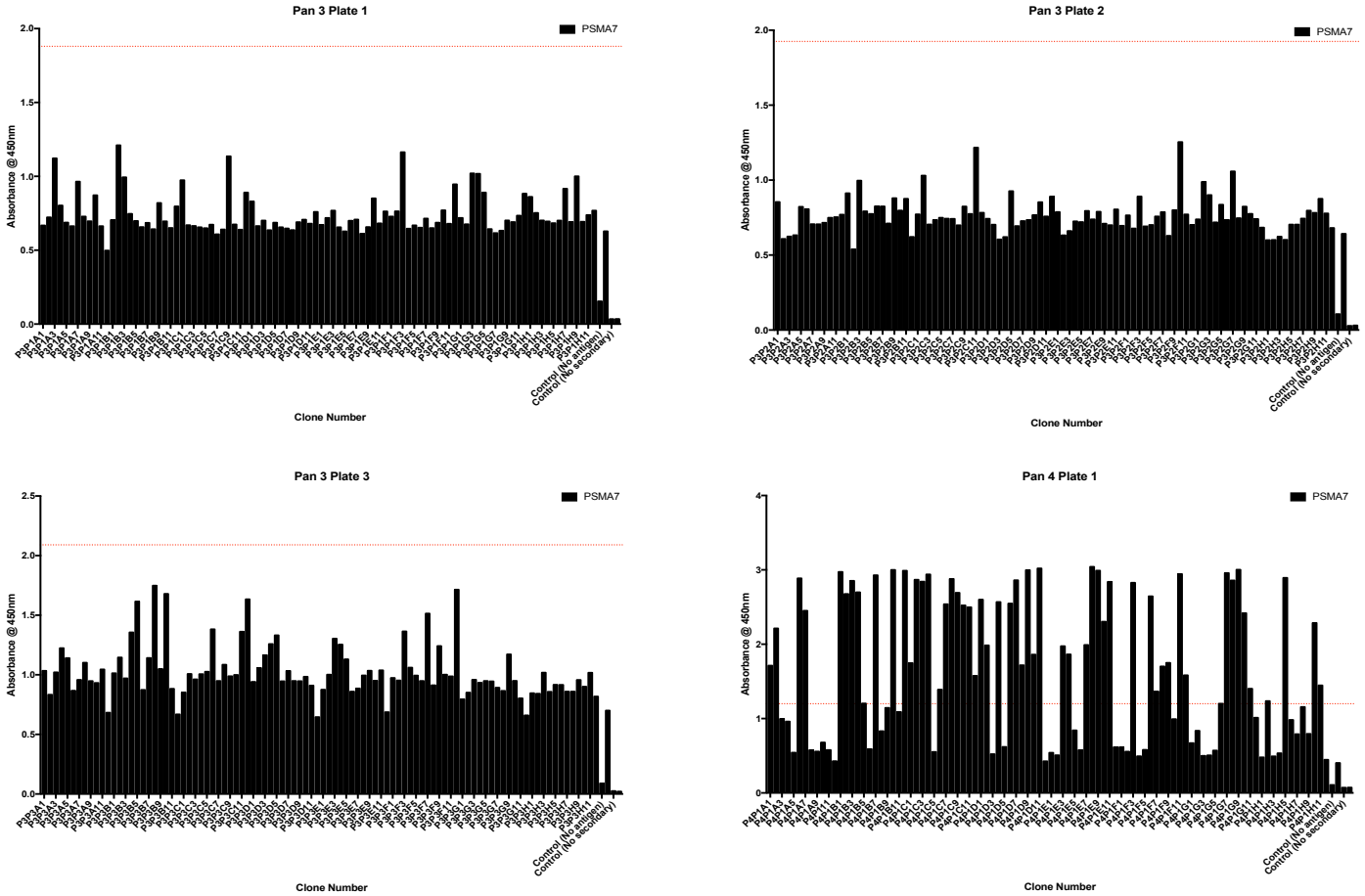

Figure 5.9 - Soluble monoclonal indirect ELISA screening for anti-PSMA7 scFv. Clones (384 in total) were selected from PCa cell panning rounds 3 and 4. The scFv-containing supernatant was incubated on PSMA7-coated plates and binding was detected using a commercial HRP-conjugated anti-HA commercial antibody and the substrate TMB. Each graph included a background cut-off that was 3 times the highest negative control value (indicated as a red horizontal line) to highlight scFv binding.

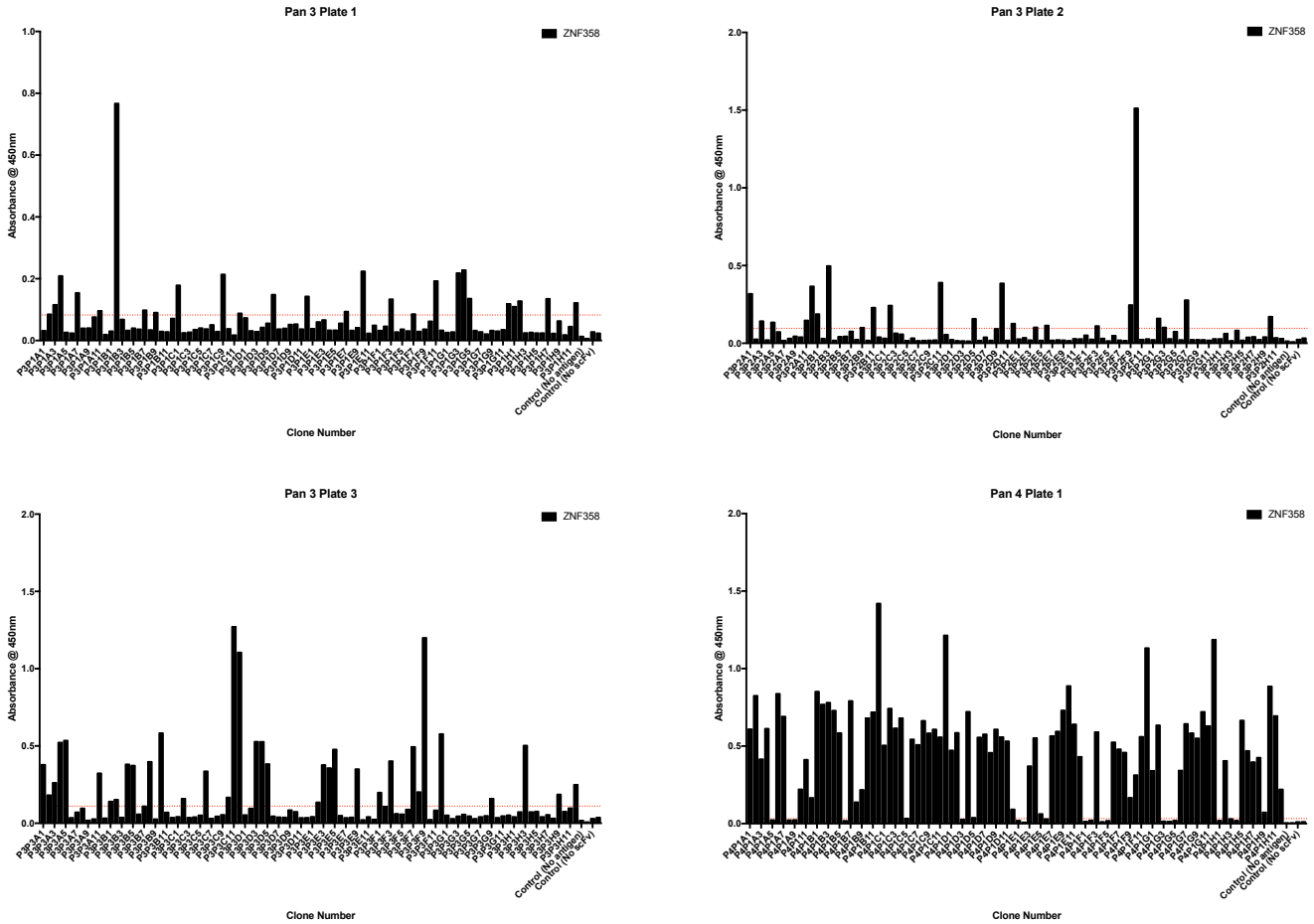

Figure 5.10 - Soluble monoclonal indirect ELISA screening for anti-ZNF358 scFv. Clones (384 in total) were selected from PCa cell panning rounds 3 and 4. The scFv-containing supernatant was incubated on ZNF358-coated plates and binding was detected using a commercial HRP-conjugated anti-HA antibody and the substrate TMB. Each graph included a background cut-off that was 3 times the highest negative control value (indicated as a red horizontal line) to highlight scFv binding.

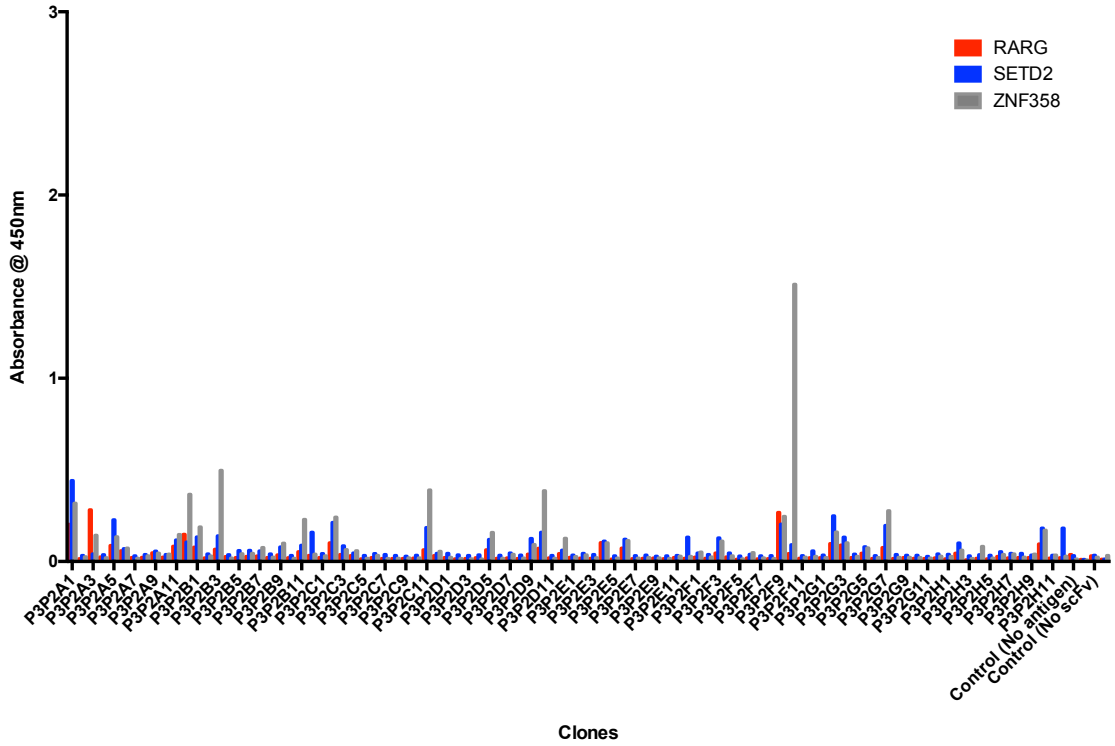

Figure 5.11 – Comparison of Round 3 monoclonal scFv, isolated in plate 2, showing binding to ZNF358, SETD2 and RARG. Clones that yielded a specific binding signal to one antigen were clones of interest. Similar comparisons were also carried out for plates 1 and 3 of panning round 3 and plate 1 of panning round 4 (data not shown).

5.2.1.4 Western blots of scFv-containing lysates of clones-of-interest

Clones displaying specific binding to one of the three antigens were selected from the monoclonal ELISAs carried out in the previous section and tested against their predicted antigens in a Western blot format (see Table 5.2) (see Section 2.5.2.3). Lysates were also analysed for the presence of scFv in a Western blot format (see Section 2.3.24). The presence of scFv in the lysates of selected anti-ZNF358 candidates P3P1B2, P3P2F10 and P3P3C11 is shown in Figure 5.12. Slight variances in the scFv size suggested genetic diversity between the clones. ScFv, while typically 28 kDa, can be as large as 36 kDa (Shen *et al.*, 2005). When the selected scFv were tested for binding to their predicted antigens, specific binding was observed from the clone P3P2F10 against ZNF358 (see Figure 5.13). The other clones analysed showed no specific binding. The lack of detectable binding may be due to the concentration of antibody in the lysate being too low or the antigen epitopes being inaccessible to the scFv in western blot format (Murphy *et al.*, 2012). It may be the case that the clones selected were not specific to their hypothesized antigens. ScFv P3P2F10 in bacterial lysate bound ZNF358 in both ELISA and Western Blot formats. Therefore, this scFv was selected to be purified for further characterisation.

Table 5.2 –Selected scFv-of-interest tested against their hypothesized antigens in Western blot format.

Clone	Antigen
P3P1B2	ZNF358
P3P1B10	SETD2
P3P1F9	SETD2
P3P2A3	RARG
P3P2B3	ZNF358
P3P2F10	ZNF358
P3P3C11	ZNF358
P3P3G12	SETD2
P4P1A3	RARG
P4P1A11	RARG
P4P1E3	RARG
P4P1E5	SETD2

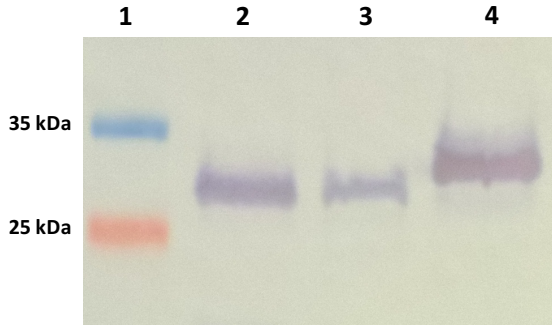

Figure 5.12 – Analysis of the crude lysates of three anti-ZNF358 candidates for the presence of scFv. Lysates (15 μL) were separated by SDS-PAGE and transferred to a membrane for Western blotting. The scFv were detected using a commercial peroxidase-conjugated anti-HA antibody and the substrate TMB. Lane 1 contained protein ladder as a size reference. Lane 2 contained lysate from the clone P3P1B2. Lane 3 contained lysate from the clone P3P2F10. Lane 4 contained lysate from the clone P3P3C11.

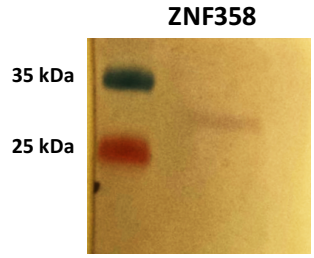

Figure 5.13 – Western blot of P3P2F10 lysate binding to ZNF358. The clone's lysate was diluted 1 in 50 in PBST containing 1% (w/v) BSA. Binding to ZNF358 (1 μg) was detected using a commercial rabbit anti-HA antibody coupled with HRP-conjugated goat anti-rabbit and the TMB chromogenic substrate.

5.2.2 Characterisation of P3P2F10 against the antigen ZNF358

5.2.2.1 Sequencing of the anti-ZNF358 candidate scFv P3P2F10

The clone P3P2F10 was sequenced to ensure that it encoded an intact scFv sequence with appropriate elements for purification (see Section 2.3.17). The retrieved sequence was compared to an scFv (P4P1E5) isolated from the fourth round of panning not considered to be specific to ZNF358, to highlight differences in the CDR regions, using the Clustal Omega alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). Following alignment, the CDR regions were identified using Kabat rules (Martin, 2015) and regions of interest were coloured to allow for comparisons (see Figure 5.14). Diversity was evident between the two clones – particularly in the CDR regions of both chains. This, in addition to intact linker, His_6 and HA sequences, suggested the scFv could be expressed and purified for characterisation.

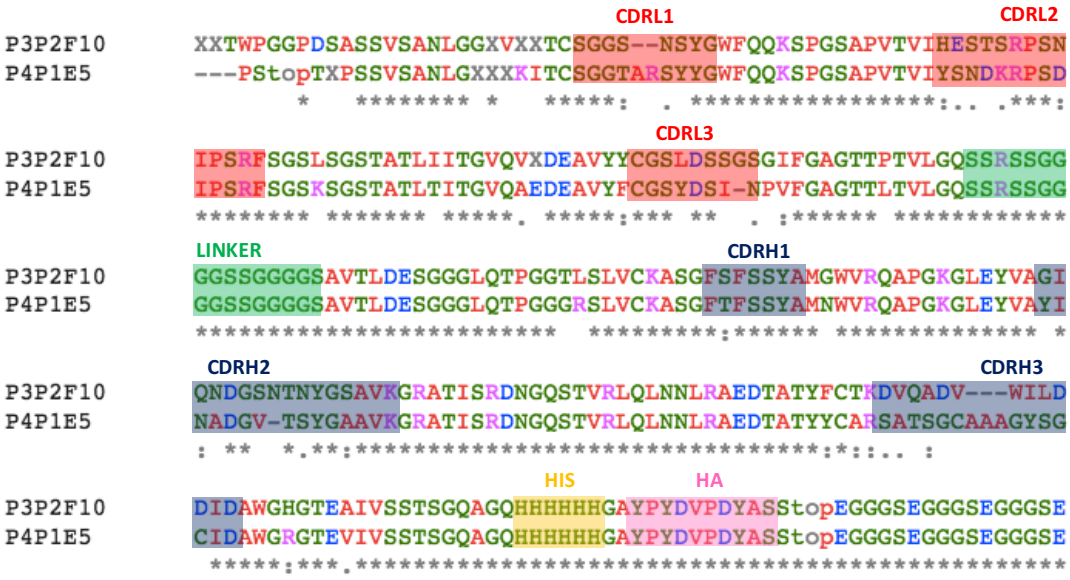

Figure 5.14- Avian anti-ZNF358 candidate P3P2F10 scFv gene sequence compared to the avian P4P1E5 scFv gene sequence for analysis using the Clustal Omega alignment tool. The highlighted sequences in red (light chain) and blue (heavy chain) show the CDRs of the recombinant antibody fragments. Also highlighted are the linker sequence (green), His_6 tag (yellow) and HA tag (pink). The intact sequence confirmed the suitability of the scFv to be purified and characterised.

5.2.2.2 Time course optimisation of P3P2F10 expression

A preliminary time-point expression optimisation was performed on the anti-ZNF358 candidate P3P2F10 prior to large-scale purification (see section 2.5.3.1). A 1 mL sample of the culture supernatant was removed every hour after IPTG induction, in addition to a sample the following morning (after 16 hours). The extracted samples were freeze thawed (x3) and stored at -20^{o}C for use in an indirect ELISA against immobilised ZNF358 (Figure 5.15). From this analysis, the overnight sample was observed to have the highest level of scFv expression.

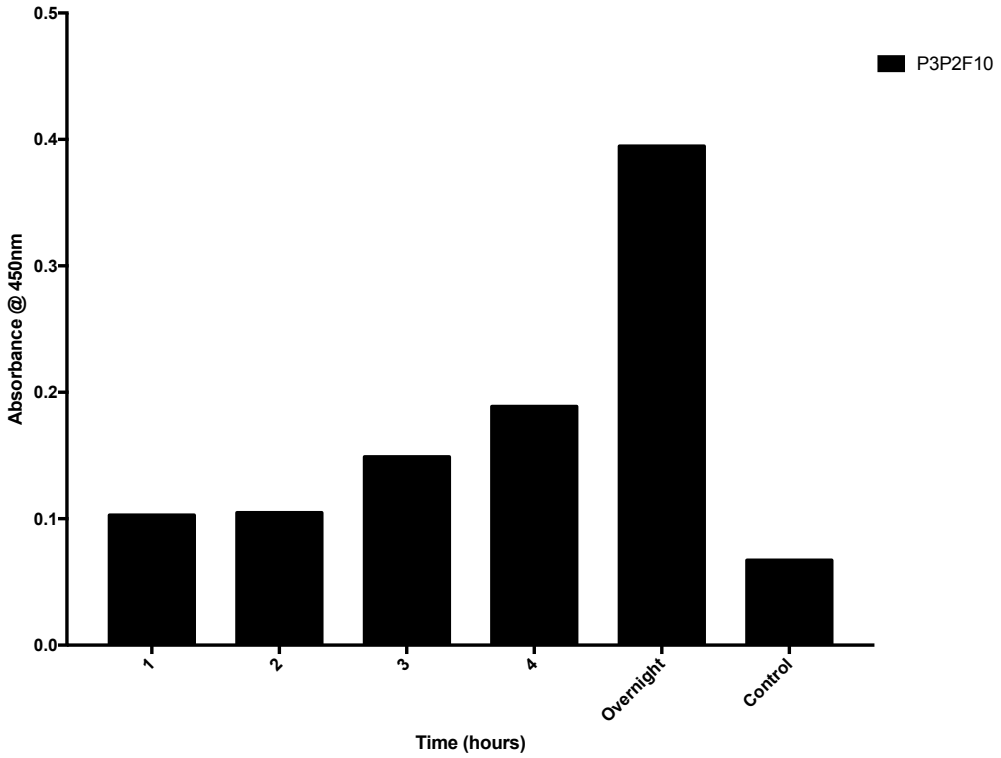

Figure 5.15 – Time-point expression optimisation of the clone P3P2F10. Culture samples were taken at hourly time points following IPTG induction (1, 2, 3, 4 hours and overnight) for analysis. The supernatant samples were diluted 1 in 5 in PBST (100 μL) and tested against immobilised ZNF358 in an ELISA. A sample lacking lysate was also included as a control. ScFv binding was detected using a commercial HRP-conjugated anti-HA antibody and developed using the substrate TMB.

5.2.2.3 Large-scale expression and IMAC purification of P3P2F10

For large-scale expression, the clone P3P2F10 was grown in a 400 mL culture and induced for expression. The scFv was released from the cells' periplasm via sonication and underwent imidazole-based IMAC purification (detailed in Section 2.5.3.2). Each step of the procedure was analysed for purity by SDS-PAGE (see Figure 5.16), and for the presence of a HA-tagged scFv by Western blotting (see 5.17). The presence of a band of the expected size (~28 kDa) in both SDS-PAGE and Western blots experiments indicated purification of the scFv. Contaminant proteins were also present in the final concentrated elution sample (hypothesized to be remnant proteins from wash step). To exclude the contaminant proteins from the calculation of protein concentration, the final scFv concentration (0.25 mg/mL) was determined using the freely available ImageJ software (Rasband, 2012) instead of a Nanodrop 1000. ImageJ compares protein band densities to known standards on an SDS-PAGE gel to achieve accurate protein quantitation (Rasband, 2012). For this quantitation, the concentrated elution fraction was analysed on an SDS-PAGE gel using BSA as a standard (see Figure 5.16 B below and Figure 8.8 in Appendix Section E).

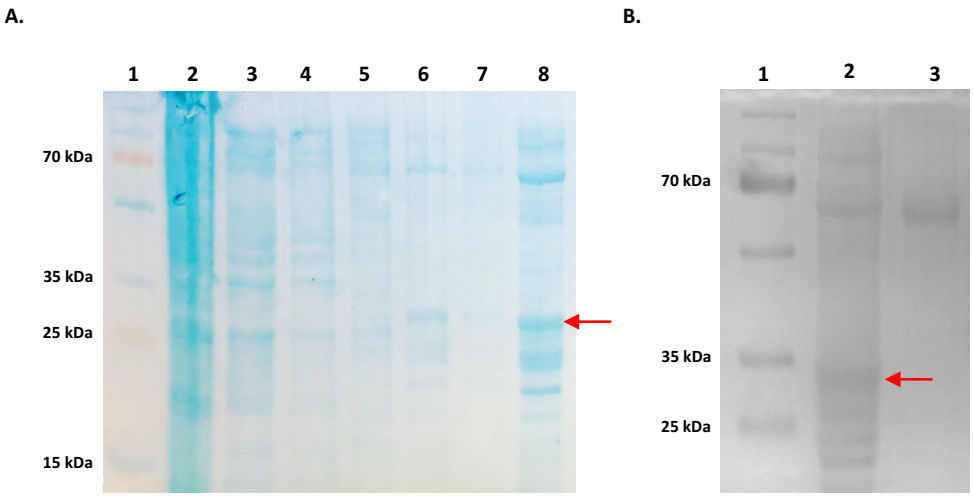

Figure 5.16 – Image A - SDS-PAGE analysis of expression and IMAC purification of the scFv P3P2F10. Lane 1 contains a protein ladder for size reference. Lane 2 contains a filtered lysate fraction. Lanes 3 and 4 contain fractions from washes A and B, respectively. Lanes 5, 6 and 7 contain samples taken from the first, middle and last mL of the eluted fraction. Lane 8 contains a sample from the final concentrated elution fraction; the scFv is highlighted using a red arrow. Image B - Comparison of P3P2F10 scFv to a BSA standard on an SDS-PAGE gel for determination of concentration using ImageJ. Lane 1 contains a protein ladder for size determination of fragments. Lane 2 contains 15µL of the purified P3P2F10 scFv, highlighted using a red arrow. Lane 3 contains 15µL of 1 mg/mL BSA. The image was converted to greyscale by the software for quantitative comparison of band intensities.

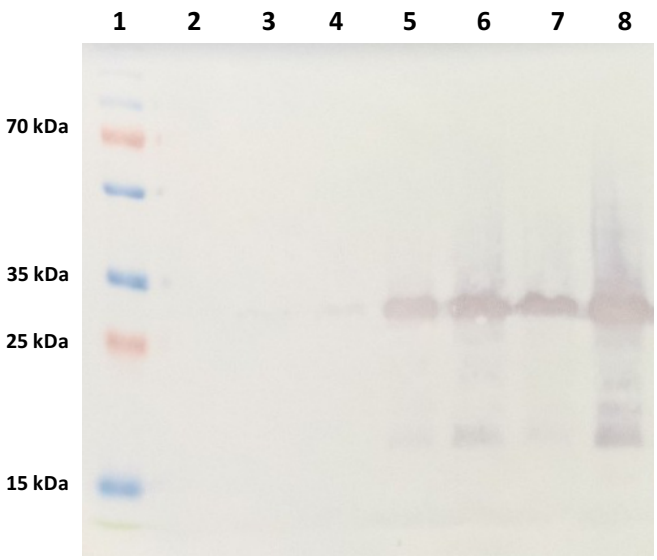

Figure 5.17 – Western blot analysis of expression and IMAC purification of the scFv P3P2F10. Lane 1 contains a protein ladder to provide a size reference. Lane 2 contains a filtered lysate fraction. Lanes 3 and 4 contain fractions from washes A and B, respectively. Lanes 5, 6 and 7 contain samples taken from the first, middle and last mL of the eluted fraction. Lane 8 contains a sample from the final concentrated eluted fraction. Binding was detected using a commercial HRP-conjugated anti-HA antibody and the TMB chromogenic substrate.

5.2.2.4 ELISA analysis of P3P2F10 binding to ZNF358

Following purification, the anti-ZNF358 candidate P3P2F10 was tested in a checkerboard ELISA experiment to investigate its binding at different dilutions (from 1 in 100 to 1 in 100,000) against varying concentrations of ZNF358-coated on plates (1 μg/mL to 10 μg/mL). Binding was detected using commercial rabbit anti-HA and HRP-conjugated goat anti-rabbit IgG secondary antibodies (see section 2.5.3.3). The results are shown in Figure 5.18 and are as expected for binding.

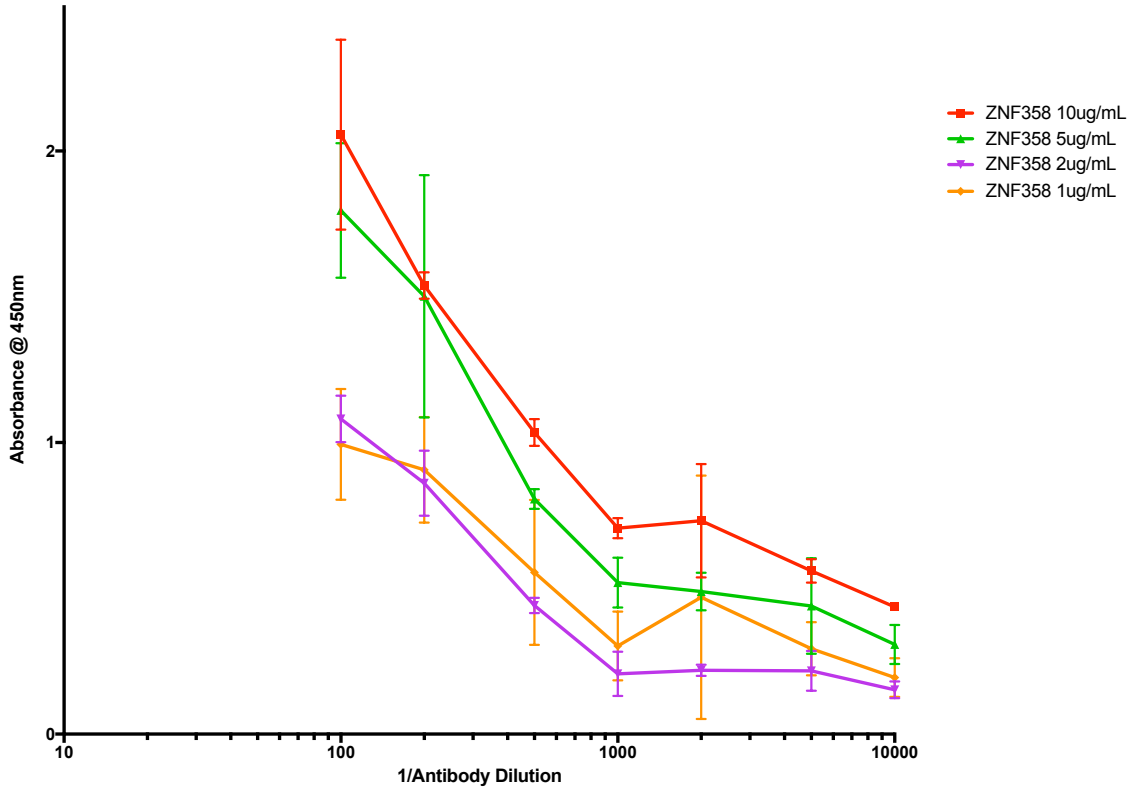

Figure 5.18 - Checkerboard indirect ELISA investigating binding between the scFv P3P2F10 and ZNF358. Varying concentrations of the antigen ZNF358 (100 µL of 1 µg/mL, 2 µg/mL, 5 µg/mL and 10 µg /mL concentrations) were coated on a 96-well plate and plate blocked to prevent non-specific binding. The anti-ZNF358 candidate P3P2F10 scFv was added to the plate at a range of dilutions (1 in 100 to 1 in 100,000). Binding was detected using commercial rabbit anti-HA and HRP-conjugated goat anti-rabbit IgG antibodies, followed by a 10-min incubation with TMB. The error bars refer to the standard deviation around the mean value of the three replicates for each condition.

5.2.2.5 Antigen specificity analysis of the P3P2F10 scFv

Following confirmation of scFv binding to ZNF358, the scFv was assessed for cross-reactivity against two protein array-positive antigens (SETD2 and RARG). A protein array-negative antigen (RASdmi) and a non-specific scAb (developed by Dr. Caroline Murphy) were also included as additional controls in the assay. The scFv P3P2F10 was tested at a 1 in 1,000 dilution against 1 μg of immobilised antigen (see section 2.5.3.4). It can be seen in Figure 5.19 that the scFv yielded a much stronger binding signal to ZNF358 than all other antigens tested, indicating specificity.

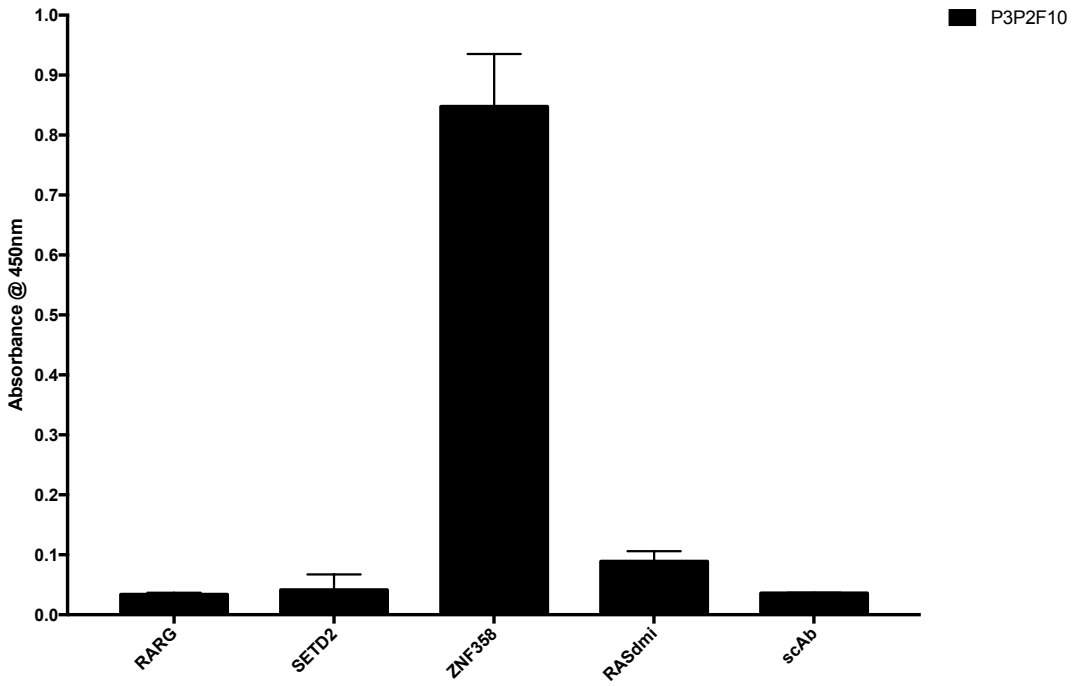

Figure 5.19 – P3P2F10 antigen cross-reactivity assay. The scFv was analysed (100 μL at a 1 in 1,000 dilution) by indirect ELISA for binding against 3 protein array-positive antigens (RARG, SETD2, ZNF358), one protein array-negative antigen (RASdmi) and one protein-array absent antigen (a non-specific scAb). Binding was detected using a commercial HRP-labelled anti-HA antibody. The bars represent the mean of three replicates for each assay condition. The error bars indicate the standard deviation.

5.2.2.6 P3P2F10 binding ZNF358 in Western blot format

The purified scFv P3P2F10 was tested against 1 μg of the antigen ZNF358 in Western blotting format (see Section 2.5.3.6). A band matching the SDS-PAGE-depicted size of the fragment was visible, indicating specific binding of ZNF358 by the scFv P3P2F10. This result, and preceding characterisation experiments of the antibody, confirmed that a scFv specific to an array-identified antigen was successfully isolated from the screening experiments.

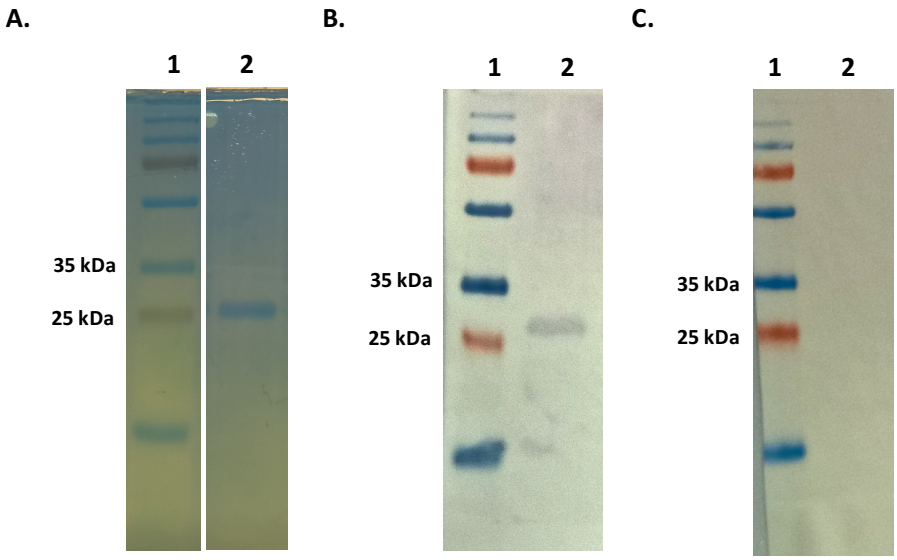

Figure 5.20 - Binding of ZNF358 by the scFv P3P2F10 in Western blot format. Lane 1 of all images (A, B and C) contains a protein ladder to provide a size reference. Lane 2 of all images contains 1 μg of ZNF358. The antigen ZNF358 (1 μg) was run on an SDS PAGE gel (shown in image A) before being transferred to a nitrocellulose membrane for Western blotting (images B and C). The scFv P3P2F10 (5 mL of a 1 in 1,000 dilution in PBST with 1% (w/v) BSA) was incubated with the membrane shown in image B and analysed for binding using commercial rabbit anti-HA and HRP-conjugated goat anti-rabbit IgG antibodies and the chromogenic substrate, TMB. The scFv was excluded from the immunoassay in image C, as a negative control.

5.2.2.7 Characterisation of P3P2F10 anti-ZNF358 binding using a Biacore 3000

The scFv P3P2F10 and ZNF358 interaction was characterised using the kinetic wizard functionality of a Biacore 3000 and the subsequent data generated were evaluated using BIAevaluation software, version 4.1.1. Biacore technology provides accurate data on ligand-analyte interactions in terms of specificity, kinetics (association and dissociation) and quantity in real-time without the need for labels (Lemass *et al.*, 2014). Here, the scFv P3P2F10 was passed over the surface of a ZNF358-immobilised chip at varying concentrations. In addition to the inclusion of buffer-only control, the antibody was also passed over a second flow cell without ZNF358 immobilised. Signals from the ZNF358-free reference surface were subtracted from the test surface signals to remove any systemic effects. The kinetic interaction is depicted in Figure 5.21 and the full kinetic profile can be viewed in Table 5.3. The K_D is the equilibrium constant and is a ratio of k_{off}/k_{on}, between the antibody and its antigen. The lower the K_D, the stronger the antibody affinity for the antigen. The K_D observed from this kinetic characterisation (2.66$\pm$1.98 nM) falls within the acceptable range of antibody-antigen affinities (Leonard *et al.*, 2007). The chi^2 value (of 0.31$\pm$0.12) indicates a good fit of the 1:1 Langmuir binding model and can be viewed in Figure 5.22. Chi^2 values derived from kinetic characterisation experiments lower than 10 are considered acceptable (Zhang and Oglesbee, 2003). Overall, the scFv shows good association and dissociation to the antigen ZNF358.

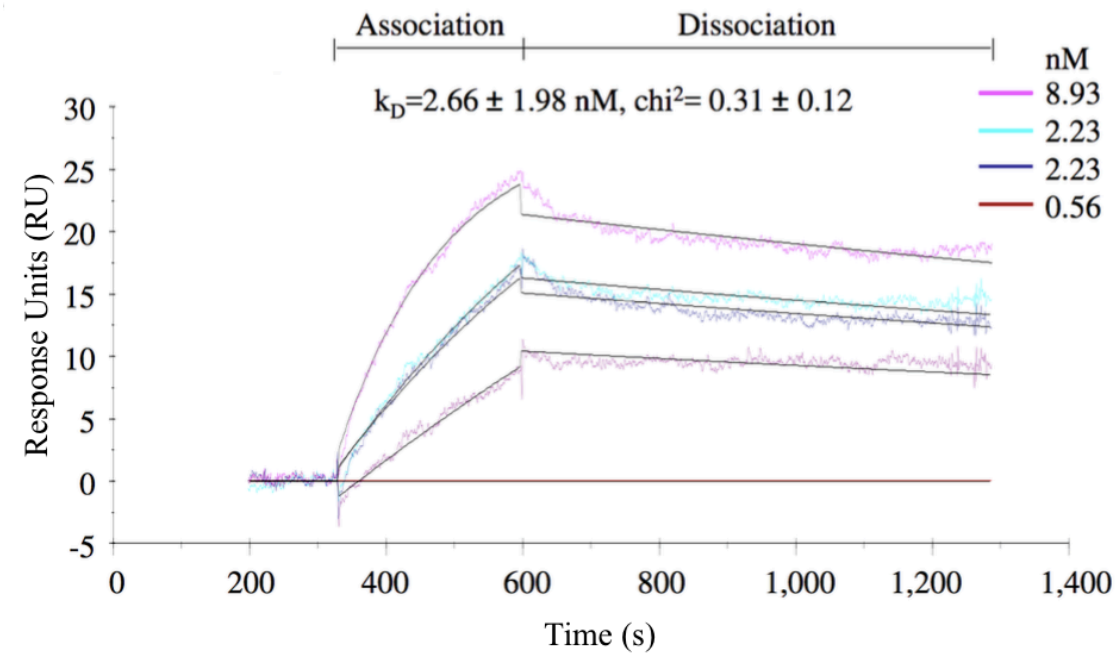

Figure 5.21 - Kinetic analysis of the interaction between ZNF358-immobilised on a CM5 Dextran chip and the scFv P3P2F10, carried out on a Biacore 3000. Four scFv concentrations were used in the analysis (8.93, 2.23, 0.56 and 0 nM) to fit the kinetics using 1:1 Langmuir binding model (global fit).

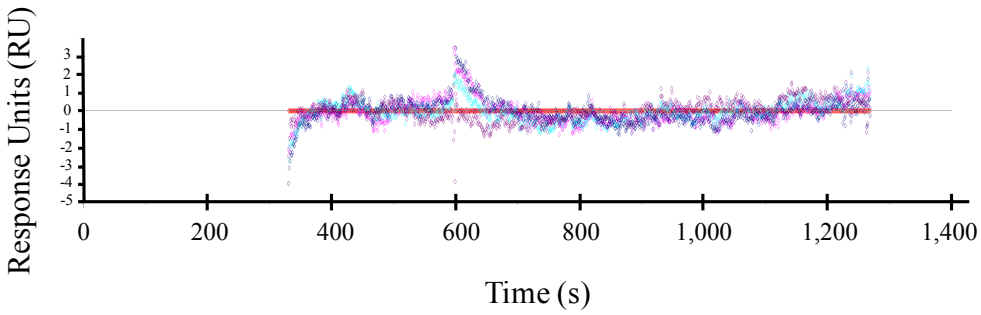

Figure 5.22 – Residual Plot showing distribution of the data points illustrating the "goodness" of the 1:1 Langmuir binding model fit of kinetic analysis depicted in Figure 5.21. A Chi² value of 0.31±0.12 was derived.

Table 5.3 – Biacore 3000-derived kinetic constants for scFv P3P2F10 binding to ZNF358 on a CM5 dextran chip.

Derived Constant	P3P2F10 scFv - ZNF358
Association Constant (K_a)	$2.82 \pm 3.5 \times 10^{5}$ $M^{-1}s^{-1}$
Dissociation Constant (K_d)	$2.93 \pm 0.13 \times 10^{-4}$ s^{-1}
Equilibrium Constant (K_D)	2.66 ± 1.98 nM
Chi^2	0.31 ± 0.12

5.3 Discussion

This chapter details the screening of 384 PCa-enriched scFv (see Chapter 3), against four selected antigens (RARG, PSMA7, SETD2 and ZNF358) identified from the protein array-profiling experiments carried out in Chapter 4. The chosen antigens varied in protein array signal intensity, with one strength 3 (SETD2), two strength 2s (PSMA7 and ZNF358) and one strength 1 (RARG). The same clones that were present on the protein arrays were acquired from the Unipex library and their identity was confirmed through analysis of their sequences. PSMA7 was found to encode the entirety of its NCBI-listed protein sequence while the ZNF358 clone encoded the central region of its listed protein sequence. The RARG and SETD2 clones were found to encode fragments at the C-terminus of their proteins. In addition, the sizes of each antigen fragment were estimated using the ExPAsy size estimation tool before expression and purification experiments. Following bioinformatic characterisation of each clone, the antigens were expressed in glucose and lactose-containing auto induction media and purified via IMAC. The resultant antigen-fragments were then investigated with SDS-PAGE and western blotting experiments. The presence of a dominant protein band that matched each clone's ExPASy-derived size-estimate, that was also bound by an anti-His_6 antibody confirmed the successful purification of each antigen.

In a preliminary screening experiment, each antigen was immobilised on 96-well ELISA plates. The 384 clones enriched against PCa from Chapter 3 were induced for scFv expression and lysed. The lysates were then analysed for binding against each antigen. Binders-of-interest were those that exhibited a strong binding signal to one antigen specifically. It was observed that a high proportion of scFv isolated from the fourth round of panning bound a multitude of antigen-coated surfaces, suggesting they were

either heterophilic antibodies or bound to BSA. PSMA7 screening yielded strong binding across all clones. This was considered to be non-specific secondary antibody binding, due to a strong binding signal observed from the negative control that excluded the scFv from the immunoassay, but not from those that excluded the antigen or secondary antibody. This indicates that PSMA7 was identified as a false positive in the protein array experiments. Thus, PSMA7 was excluded from subsequent screening assays. The monoclonal ELISAs revealed multiple potential anti-ZNF358 candidate scFv; however, specific binders to SETD2 and RARG were less obvious. For future work, this primary screening experiment may be improved upon by modifying the immunoassay format. An alternative format could be the immobilisation of scFv (in lysate) to the plate surface and probing with fluorophore-coated antigens. This approach would simplify the assay - removing antibodies that bind to BSA as false positives, as well as removing the requirement for secondary antibodies, hence reducing the risk of non-specific binding. A more sensitive approach would be to immobilise multiple antigens to the chip of a high throughput, SPR-based platform, and passing monoclonal scFv-containing lysate over the surface. Commercially available platforms capable of carrying out such analyses include the Biacore 4000 (http://proteins.gelifesciences.com/Products-for-proteins/SPR-systems/Biacore-4000), which operates in a similar manner to the Biacore 3000 but can perform up to 16 kinetic assays in parallel, and the IBIS-MX96 (http://www.ibis-spr.nl/product/ibis-mx96/), which utilises SPR imaging to simultaneously analyse up to 96 binding interactions. Such platforms would allow the user to screen each clone, label-free, against multiple antigens simultaneously and provide real-time antibody characterisation of an increased sensitivity (Leonard *et al.*, 2007). It was hoped to use a Biacore 4000 to do this but unfortunately it was damaged and not available.

A total of 12 scFv were selected from the preliminary ELISA screening experiments (four candidate scFv to each antigen). The crude lysate of each selected clone was then tested against its hypothesized antigen in western blot format. Of the scFv selected, only one clone (P3P2F10) was observed to specifically bind its target antigen. The clones that did not bind their antigens may not have been specific or may have required further optimisation of the experimental conditions (such as the purity and concentration of the scFv, as well as the incubation times and temperatures). Another possibility is that the antigens were displayed differently between the protein arrays and the Western blots or ELISA screening assays and thus not bound (Murphy *et al.*, 2012). The clone P3P2F10 was carried forward for further characterisation as it exhibited selective binding in lysate to ZNF358 both in ELISA and Western blot formats.

Sequencing of P3P2F10 revealed a complete scFv sequence, with HA and His_6 tags, an intact linker and unique CDRs (when compared with the sequence of another scFv clone isolated in chapter 3). Following a small-scale optimisation of expression experiment, the P3P2F10 scFv was expressed in a 400 mL culture and purified by IMAC. SDS-PAGE and Western blot analysis of the purified product revealed the presence of a scFv fragment. Instead of using the BCA assay or a Nanodrop 1000, both of which quantify the total protein content in a solution, the scFv was quantified by comparing its band intensity with known BSA standards on and SDS-PAGE gel using the National Institute of Health-sourced ImageJ software (Rasband, 2012). This quantitation approach was utilised so the contaminant protein bands, that were non-specifically purified along with the scFv, could be excluded. The purified P3P2F10 scFv was then shown to bind ZNF358, both in Western blot and ELISA formats. A cross-reactivity investigation was also carried

out where the scFv's binding to ZNF358 was compared to the other two selected protein array-identified antigens (SETD2 and RARG), as well as an array present protein (RASdmi) that was not identified as positive in the protein array experiments, and a non-specific scAb that was not present on the protein array. This experiment yielded a scFv binding signal to ZNF358 over ten times stronger than the signal observed from the array-positive antigens SETD2 and RARG, confirming its specificity. Finally, the scFv's binding to ZNF358 was characterised using a Biacore 3000. The antigen was immobilised to the surface of a CM5 dextran chip and varying concentrations of the P3P2F10 scFv were passed over the surface. The K_D of the interaction was determined to be 2.66$\pm$1.98 nM, demonstrating the successful isolation of a high-affinity scFv to ZNF358.

This scFv screening strategy demonstrated in this chapter would not have been achievable without the preceding protein array-profiling experiments carried out in Chapter 4. This array-coupled strategy improves upon the current screening approaches, by providing the user with a list of antigens corresponding to an isolated scFv panel and allowing the antigen profile to be assessed for value before starting screening. Without prior protein array profiling, a 'blind' screening strategy would have been required. For example, following the isolation of 994 clones from a cell-panned, naive scFv library, Keller *et al.* (2015) sequenced all clones and removed any duplicates (391) and clones yielding incomplete sequencing data (437). The resultant 166 scFv were expressed in lysate and screened for binding to target cell surfaces in a cell ELISA. A total of 86 clones were selected as high affinity binders, from which, 8 scFv underwent detailed characterisation. These 8 scFv were purified and characterised further by cell ELISAs and fluorescent microscopy, from which a single antibody was selected based on its binding properties. The selected scFv was immunoprecipitated with cell lysate to

allow for its corresponding antigen to be identified via mass spectrometry. This antibody binding characterisation, followed by antigen identification by mass spectrometry, or 'function-before-identity', is the current standard approach to generating recombinant antibodies against non-predetermined antigens (Siva *et al.*, 2008; Even-Desrumeaux *et al.*, 2013). While useful for isolating high-affinity binders to surface antigens expressed on a cell line of interest, the strategy is both resource and labour-intensive. Additionally, it focuses on isolating high affinity binders, yet affinity is not always the best way to screen for therapeutics, as low affinity binders have been shown to better penetrate tumours (Rudnick *et al.*, 2011).

Interestingly, the successfully screened anti-ZNF358 scFv characterised and validated in this chapter, was profiled as a moderate binder during the protein array experiments. While an antibody was to SETD2, (characterised as a strong binding signal on the arrays), was not identified. As such, the major bottle-neck observed in the strategy developed in this research appears to be down-stream screening for antigens/scFv pairs, following array-identification. This may not a major issue, provided multiple identified antigens are considered to be of interest. Here, screening for four antigens was carried out, with an scFv to only one being isolated. To increase the output of scFv generation and characterisation, improved methods of post-protein array characterisation need to be implemented. For instance, Biacore 4000 array methods allow for rapid and more sensitive characterisation of dozens of antigens to several hundred scFv. The isolation of weak-binding scFv to antigens-of-interest would still be of value, as recombinant antibody technology allows for the affinity of an antibody-antigen interaction to be improved through post-characterisation modifications/engineering, if desired

(Fitzgerald *et al.*, 2011). The current standard 'function-before-identity' screening approach risks removing these low-affinity binders at an early stage of the process.

In summary, this chapter has reversed the standard 'function-before-identity' antibody generation strategy to generate antibodies to cancer relevant targets that have not been pre-determined. Revealing the antigen profile of an antibody panel, before screening for binders-of-interest drastically decreases the cost and labour required to isolate specific antibodies of value. Adoption of this strategy in the field may increase the generation of antibodies to antigens of as yet unknown importance in the cancer-disease state – paving the way for the discovery of novel diagnostic and therapeutic targets.

Chapter 6 – Overall Conclusions

Antibody technology has made vast improvements to the detection and treatment of many cancers; however, it is limited in the field of PCa due to a lack of biomarker targets. Recombinant antibody technology has also been established as a viable approach to generating antibodies to unknown antigens present in cancer cells. The technology has the capability to discover biomarkers and simultaneously produce a specific antibody to that target. The major drawback of the technology, however, is the requirement for extensive antibody characterisation experiments to reveal the identity of each antigen corresponding to the generated antibodies. Thus, a common theme in the field is to carry out stringent, function-focused antibody characterisation experiments and subsequently identify a subset of the antigens to the antibodies. The number of antibody binders selected for mass spectrometry-based antigen identification following screening varies by publication, but rarely exceeds ten (Siva *et al.*, 2008; Even-Desrumeaux *et al.*, 2013; Keller *et al.*, 2015) and in some cases is omitted entirely, where the isolated antibodies are characterised based on their phenotype and their antigens remain unknown (Popkov *et al.*, 2004; Bowes *et al.*, 2011;). This reliance on antibody selection prior to antigen identification risks overlooking antibodies to antigens of value and is a major limiting factor in the production of antibodies to non-predetermined antigens. Providing a solution to this shortcoming would greatly increase the rate at which antibodies can be produced to targets that may be implicated in PCa.

The principal aim of this research was to provide these improvements to develop recombinant antibodies to non-predetermined, cancer-associated antigens. The strategy applied in this research was to generate large non-predetermined antibody libraries directed towards PCa, through live cell chicken immunisation and panning.

Protein array technology was exploited to reveal the antigen repertoire at each stage in the generation process and negate the current requirement of 'function-before-identity' antibody screening.

Chapter 3 provides an overview of the generation of a recombinant antibody library against antigens present in PCa cells. As expected, the immunisation of two avian hosts (chickens ABG048 and ABG055) with PCa cell lines successfully induced an anti-PCa response. Conversion of these immune responses into scFv-phage libraries resulted in a discrepancy in sizes by a factor of 300, with the library derived from the chicken ABG055 being the larger (5×10^8 c.f.u). Thus, the ABG055 scFv-phage library was more likely to contain scFv-of-interest and was selected for anti-PCa enrichment. PCa cells were used to pan the scFv-phage pool to produce a diverse set of antibodies enriched against antigens present in PCa. Enrichment of the library against PCa cells was confirmed by polyclonal scFv-phage ELISA and a total of 384 clones were isolated.

Using protein array technology in Chapter 4, it was possible to characterise the antigen profile at each stage in the antibody generation process carried out in Chapter 3. Prior to use, commercial polyclonal secondary antibodies selected to be applied in pairs for IgY, scFv-phage and scFv-profiling experiments were profiled in the absence of their intended antigens as a control experiment, characterising their specificity. This allowed for the identification and removal of false-positive signals from the antigen lists generated downstream, as well as providing a reference list of cross-reactive proteins for the antibodies should they be used in future experiments. The importance of such secondary antibody validation experiments, before use in complex immunoassays, is

now being emphasized in the scientific community (Lemass *et al.*, 2016; Uhlen *et al.*, 2016).

Antigen profiling of sera-contained IgY, isolated from the chicken ABG055 before and after the PCa immunisation regime, allowed for the avian immune response to PCa cell lines to be revealed. An improved understanding of how an avian host's immune system responds to human proteins present in the cells adds value to future antibody generation projects. *In silico* analysis of the immune response, as expected, showed no enrichment based on cellular processes, locations, pathways or functions. This analysis did however demonstrate a practical way to interpret the large datasets produced by protein arrays - grouping antigens based on their known functions, cellular locations or involvement in biochemical pathways. Generating antibodies to antigens implicated in biological processes relevant to cancer, such as cell cycle regulation, may identify novel PCa biomarkers and therapeutic targets. Despite no enrichment in the immune response to antigen phenotype being observed, analyses comparing the homology of array-identified antigens to their orthologs in chicken revealed a trend towards highly conserved orthologs. Analysis of the protein array gene set likewise displayed an enrichment towards conserved orthologs and was considered to be a contributing factor, although the trend in enrichment towards conserved orthologs was still stronger in the immune response dataset. This finding conflicts with the general consensus in the field, that chickens will produce a robust immune response against mammalian proteins that are not homologous to the native avian proteins (Bowes *et al.*, 2011). For future work, observing a similar conserved-ortholog enriched immune response, using protein arrays that better match the human genome in terms of ortholog conservation, would strengthen these novel findings.

Protein array technology was also applied to characterise the scFv-phage library produced from the chicken's immune response, as well as the 384 antibodies isolated following panning. A high background signal was observed in the scFv-phage profiling, making it difficult to identify antigen binding signals. This was most likely caused by non-specific binding of phage to the array. From this, it was concluded that phage-displayed antibodies are not well suited for antigen characterisation using this strategy. The 384 scFv on the other hand exhibited strong array-binding signals with minimal background, producing a dataset of 166 unique antigens. The protein arrays used in this research had previously been shown to be suitable for characterising low quantities of recombinant antibodies in parallel (Holt *et al.*, 2000). This study demonstrated for the first time, the ability of the same arrays to profile high numbers of recombinant antibodies in parallel. Each high through-put antigen characterisation experiment in Chapter 4 was achieved using a single set of protein arrays, offering improved cost, speed and output data compared to the current stringent screening and mass spectrometry-reliant methods. The technology provided antigen profiles from the start to finish of the antibody generation process and antigen overlap was observed between linked datasets. Being able to monitor the antibody generation process as early as the immune response phase can direct the researcher towards antigens that may be relevant in PCa. In this research, the antibody library was enriched towards antigens present in PCa cells rather than individual antigens. However, antibodies to non-predetermined antigens may be selected as early as the post-immunisation phase by panning the library with the array-identified antigen selected from the immune response. The most valuable discovery of the research presented here is single-step antigen profiling of the scFv enriched towards PCa cells, as it provides a viable alternative to the stringent screening and mass

spectrometry-reliant approaches currently the standard in the field. Revealing the antigen repertoire before screening reversed the 'function-before-identity' approach currently used in the field. This allowed for the antigens of the scFv to be evaluated before screening for binders-of-interest.

Following the antigen profiling, the scFv were screened against a subset of selected antigens using standard immunoassay techniques. Antibodies to the antigens PSMA7, SETD2, RARG and ZNF358 were considered to be of value, as their cognate antigens had been implicated in PCa, or remain uncharacterised. The antigens were acquired from the same recombinant clones present on the protein arrays and applied to select scFv using ELISA and Western blotting. From these experiments, it was observed that the secondary antibody bound non-specifically to PSMA7, which was subsequently considered to be a false-positive. Comparisons of scFv (in crude lysate) binding to the 3 remaining antigens identified several candidate scFv that appeared specific to one antigen. However, of the scFv selected, only the clone P3P2F10 was found to bind its hypothesized antigen, ZNF358, both in ELISA and Western blot formats. This difficulty in identifying specific antibody to the other two antigens was a bottleneck in the screening carried out. In future work, this screening strategy may be improved by utilising a high sensitivity screening platform, such as a Biacore 4000, to accurately rank antibody-antigen interactions (Leonard *et al.,* 2007). Purification and further characterisation experiments of the scFv from the clone P3P2F10 demonstrated that the scFv bound specifically to ZNF358 in an ELISA format. Finally, this binding interaction was kinetically characterised using a Biacore 3000 instrument to conclude that an anti-ZNF358 antibody had been successfully screened.

Overall, this body of research demonstrated a novel approach to generating antibodies to antigens present in cancer cells, without the need for antigen preselection or downstream antigen identification using MS. Applying protein array technology to elucidate the antigen profile at each stage of the generation process allowed for the user to look inside the 'black box' of recombinant antibody production and select antigens that antibodies could successfully be generated against. Reversal of the 'function-before-identity' antibody generation strategy allowed for the antigens to be evaluated before proceeding with downstream antibody experiments. The scFv isolated as a result of this approach is the first non-polyclonal antibody generated to ZNF358 (Antibodypedia, 2016) and may be utilised in the future studies to investigate the relevance of ZNF358 in PCa. It is hoped that adoption of this novel 'identity-before-function' strategy will see an increase the output of antibodies to antigens that may otherwise be overlooked using the current-standard approaches. Such novel antibodies will improve our understanding of the disease and may be applicable to improving the current PCa detection methods or therapeutic strategies.

Chapter 7 – Bibliography

Abdiche, Y. N., Miles, A., Eckman, J., Foletti, D., Van Blarcom, T. J., Yeung, Y. A., Pons, J. and Rajpal, A. (2014). High-throughput epitope binning assays on label-free array-based biosensors can yield exquisite epitope discrimination that facilitates the selection of monoclonal antibodies with functional activity. *PLoS ONE*, 9 (3), e92451.

Andriole, G. L., Bostwick, D. G., Otis W. Brawley, Gomella, L. G., Marberger, M., Montorsi, F., Pettaway, C. A., Tammela, T. L., Teloken, C., Tindall, D. J., Somerville, M. C., Wilson, T. H., Fowler, I. L. and Rittmaster, R. S. (2010). Effect of dutasteride on the risk of prostate cancer. *European Urology*, 362, 1192–1202.

Andris-Widhopf, J., Rader, C., Steinberger, P., Fuller, R. and Barbas Iii, C. F. (2000). Methods for the generation of chicken monoclonal antibody fragments by phage display. *Journal of Immunological Methods*, 242, 159–181.

Antibodypedia (2016). *ZNF358* [Online]. Available from: https://www.antibodypedia.com/gene/12050/ZNF358 [Accessed 1 November 2016].

Bansal, D., Undela, K., D'Cruz, S. and Schifano, F. (2012). Statin Use and Risk of Prostate Cancer: A Meta-Analysis of Observational Studies. *PLoS ONE*, 7 (10), e46691.

Barbas III, C. F., Burton, D. R., Scott, J. K. and Silverman, G. J. (2001). *Phage Display: A laboratory manual*, first edition.Cold Spring Harbour Laboratory Press, New York, USA.

Berthold, D. R., Pond, G. R., Soban, F., De Wit, R., Eisenberger, M. and Tannock, I. F. (2008). Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer: Updated survival in the TAX 327 study. *Journal of Clinical Oncology*, 26 (2), 242–245.

Bowes, T., Hanley, S. A., Liew, A., Eglon, M., Mashayekhi, K., O'Kennedy, R., Barry, F., Taylor, W. R., O'Brien, T., Griffin, M. D., Finlay, W. J. and Greiser, U. (2011). Developing Cell-Specific Antibodies to Endothelial Progenitor Cells Using Avian Immune Phage Display Technology. *Journal of Biomolecular Screening*, 16 (7), 744–754.

Brezina, S., Soldo, R., Kreuzhuber, R., Hofer, P., Gsur, A. and Weinhaeusel, A. (2015). Immune-Signatures for Lung Cancer Diagnostics: Evaluation of Protein Microarray Data Normalization Strategies. *Microarrays*, 4 (2), 162–187.

Brimo, F. and Epstein, J. I. (2012). Immunohistochemical pitfalls in prostate pathology. *Human Pathology* 43(3), 313-324.

Brum, I. J. B., Martins-de-Souza, D., Smolka, M. B., Novello, J. C. and Galembeck, E. (2009). Web Based Theoretical Protein pI, MW and 2DE Map. *Journal of Computer Science Systems Biology*, 2 (1), 93–96, doi:10.4172/jcsb.1000020. Available from: http://www.omicsonline.com/ArchiveJCSB/2009/February/03/JCSB2.93.php [Accessed 30 November 2016]

Büssow, K., Cahill, D., Nietfeld, W., Bancroft, D., Scherzinger, E., Lehrach, H. and Walter, G. (1998). A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Research*, 26 (21), 5007–5008.

Büssow, K., Nordhoff, E., Lübbert, C., Lehrach, H. and Walter, G. (2000). A human cDNA library for high-throughput protein expression screening. *Genomics*, 65 (1), 1–8.

Byrne, H., Conroy, P. J., Whisstock, J. C. and O'Kennedy, R. J. (2013). A tale of two specificities: bispecific antibodies for therapeutic and diagnostic applications. *Trends in Biotechnology*, 31 (11), 621–632.

Carson, C. and Rittmaster, R. (2003). The role of dihydrotestosterone in benign prostatic hyperplasia. *Urology*, 61(4), 2–7

Cervino, C., Weber, E., Knopp, D. and Niessner, R. (2008). Comparison of hybridoma screening methods for the efficient detection of high-affinity hapten-specific monoclonal antibodies. *Journal of Immunological Methods*, 329 (1–2), 184–193.

Chen, L., Stacewicz-, M., Duncan, C., Sharifi, R., Ghosh, L., Breemen, R. Van, Ashton, D. and Bowen, P. E. (2001). Oxidative DNA Damage in Prostate Cancer Patients Consuming

Tomato Sauce-Based Entrees as a Whole-Food Intervention. *Journal of the National Cancer Institute*, 93 (24), 1872-1879.

Chen, M., Zhu, N., Liu, X., Laurent, B., Tang, Z., Eng, R., Shi, Y., Armstrong, S. A. and Roeder, R. G. (2015) JMJD1C is required for the survival of acute myeloid leukemia by functioning as a coactivator for key transcription factors. *Genes and Development*, 29 (20), 2123–2139.

Conroy, P. J., Hearty, S., Leonard, P. and O'Kennedy, R. J. (2009). Antibody production, design and use for biosensor-based applications. *Seminars in Cell and Developmental Biology*, 20, 10–26.

Conroy, P. J., Law, R. H. P., Gilgunn, S., Hearty, S., Caradoc-Davies, T. T., Lloyd, G., O'Kennedy, R. J. and Whisstock, J. C. (2014). Reconciling the structural attributes of avian antibodies. *The Journal of Biological Chemistry*, 289 (22), 15384–15392.

Conroy, P. J., O'Kennedy, R. J. and Hearty, S. (2012). Cardiac troponin I: a case study in rational antibody design for human diagnostics. *Protein Engineering, Design & Selection : PEDS*, 25 (6), 295–305.

Cook, L. S., Goldoft, M., Swartz, S. M. and Weiss, N. S. (1999). Incidence of adenocarcinoma of the prostate in Asian immigrants to the United States and their decendants. *The Journal of Urology*, 161 (1), 152–155.

Derda, R., Tang, S. K. Y., Li, S. C., Ng, S., Matochko, W. and Jafari, M. R. (2011). Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules*, 16 (2), 1776–1803.

Dias da Silva, W. and Tambourgi, D. V. (2010). IgY: a promising antibody for use in immunodiagnostic and in immunotherapy. *Veterinary Immunology and Immunopathology*, 135 (3–4), 173–180.

Drucker, E. and Krapfenbauer, K. (2013). Pitfalls and limitations in translation from

biomarker discovery to clinical utility in predictive and personalised medicine. The *EPMA Journal*, 4 (1).

El-Haibi, C. P., Singh, R., Gupta, P., Sharma, P. K., Greenleaf, K. N., Singh, S., Lillard, J. W. and Jr. (2012). Antibody Microarray Analysis of Signaling Networks Regulated by Cxcl13 and Cxcr5 in Prostate Cancer. *Journal of Proteomics & Bioinformatics*, 5 (8), 177–184.

Even-Desrumeaux, K., Nevoltris, D., Lavaut, M. N., Alim, K., Borg, J.-P., Audebert, S., Kerfelec, B., Baty, D. and Chames, P. (2013). Masked Selection: A Straightforward and Flexible Approach for the Selection of Binders Against Specific Epitopes and Differentially Expressed Proteins by Phage Display. *Molecular & Cellular Proteomics*, 13(2), 653–665.

Finlay, W. J. J., Bloom, L. and Cunningham, O. (2011). Optimized generation of high-affinity, high-specificity single-chain Fv antibodies from multiantigen immunized chickens. *Methods in Molecular Biology (Clifton, N.J.)*, 681, 383–401.

Finlay, W. J. J., DeVore, N. C., Dobrovolskaia, E. N., Gam, A., Goodyear, C. S. and Slater, J. E. (2005). Exploiting the avian immunoglobulin system to simplify the generation of recombinant antibodies to allergenic proteins. *Clinical and Experimental Allergy : Journal of the British Society for Allergy and Clinical Immunology*, 35 (8), 1040–1048.

Finlay, W. J. J., Shaw, I., Reilly, J. P. and Kane, M. (2006). Generation of high-affinity chicken single-chain Fv antibody fragments for measurement of the *Pseudonitzschia pungens* toxin domoic acid. *Applied and Environmental Microbiology*, 72 (5), 3343–3349.

Fitzgerald, J., Leonard, P., Darcy, E., Danaher, M. and O'Kennedy, R. (2011). Light-chain shuffling from an antigen-biased phage pool allows 185-fold improvement of an anti-halofuginone single-chain variable fragment. *Analytical Biochemistry*, 410 (1), 27–33.

Fitzgerald, S., Sheehan, K. M., Espina, V., O'Grady, A., Cummins, R., Kenny, D., Liotta, L., O'kennedy, R., Kay, E.W and Kijanka G. S. (2015a). High CerS5 expression levels associate with reduced patient survival and transition from apoptotic to autophagy signalling

pathways in colorectal cancer. *The Journal of Pathology: Clinical Research*, 1 (1), 54–65.

Fitzgerald, S., Sheehan, K. M., O'Grady, A., Kenny, D., O'Kennedy, R., Kay, E. W. and Kijanka, G. S. (2013). Relationship between epithelial and stromal TRIM28 expression predicts survival in colorectal cancer patients. *Journal of Gastroenterology and Hepatology*, 28 (6), 967–974.

Fitzgerald, V., Manning, B., Donnell, B. O., Reilly, B. O., Sullivan, D. O., Kennedy, R. O. and Leonard, P. (2015b). Exploiting highly ordered subnanoliter volume microcapillaries as microtools for the analysis of antibody producing cells. *Analytical Chemistry*, 87 (2), 997-1003.

Fizazi, K., Carducci, M., Smith, M., Damião, R., Brown, J., Karsh, L., Milecki, P., Shore, N., Rader, M., Wang, H., Jiang, Q., Tardos, S., Dansey, R. and Goessl, C. (2011). Denosumab versus zoledronic acid for treatment of bone metastases in men with castration-resistant prostate cancer: a randomised, double-blind study. *Lancet*, 377 (9768), 813–822.

FujiFilm (2016). *Fla-5100 Information sheet* [online]. Available from: http://www.raytek.co.uk/Brochures/FLA5100 UK-IE Brochure.pdf [Accessed 16 November 2016].

Gann, P. H., Ma, J., Giovannucci, E., Willett, W., Sacks, F. M., Hennekens, C. H. and Stampfer, M. J. (1999). Lower prostate cancer risk in men with elevated plasma lycopene levels: results of a prospective analysis. *Cancer Research*, 59 (6), 1225–1230.

Genentech (2016). Herceptin® (trastuzumab) Development Timeline [online]. Available from: https://www.gene.com/media/product-information/herceptin-development-timeline [Accessed 30 November 2016].

Gilgunn, S., Conroy, P. J., Saldova, R., Rudd, P. M. and O'Kennedy, R. J. (2013). Aberrant PSA glycosylation--a sweet predictor of prostate cancer. *Nature Reviews. Urology*, 10 (2), 99–107.

Glassman, P. M. and Balthasar, J. P. (2014). Mechanistic considerations for the use of monoclonal antibodies for cancer therapy. *Cancer Biology & Medicine*, 11 (1), 20–33.

Haas, G. P., Delongchamps, N., Brawley, O. W., Wang, C. Y. and de la Roza, G. (2008). The worldwide epidemiology of prostate cancer: perspectives from autopsy studies. *The Canadian Journal of Urology*, 15 (1), 3866–3871.

Harris, W. P., Mostaghel, E., Nelson, P. S. and Montgomery, B. (2009). Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nature Clinical Practice Urology*, 6 (2), 76–85.

Heidenreich, A., Bastian, P. J., Bellmunt, J., Bolla, M., Joniau, S., van der Kwast, T., Mason, M., Mateev, V., Wiegel, T. and Mottet, N. (2014a). EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *European Urology*, 65, 124–137.

Heidenreich, A., Bastian, P. J., Bellmunt, J., Bolla, M., Joniau, S., van der Kwast, T., Mason, M., Mateev, V., Wiegel, T. and Mottet, N. (2014b). EAU guidelines on prostate cancer. Part II: Treatment of advanced, relapsing, and castration-resistant prostate cancer. *European Urology*, 65, 467–479.

Holt, L. J., Büssow, K., Walter, G. and Tomlinson, I. M. (2000). By-passing selection: direct screening for antibody-antigen interactions using protein arrays. *Nucleic Acids Research*, 28 (15), E72.

Hommelsheim, C. M., Frantzeskakis, L., Huang, M. and Ülker, B. (2014). PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Scientific Reports*, 4, 5052.

Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. *Nature Biotechnology*, 23 (9), 1105–1116.

Hoogenboom, H. R., Lutgerink, J. T., Pelsers, M. M. a L., Rousch, M. J. M. M., Coote, J., Van Neer, N., De Bruine, A., Van Nieuwenhoven, F.A., Glatz, J.F. and Arends, J.W. (1999). Selection-dominant and nonaccessible epitopes on cell-surface receptors revealed by cell-panning with a large phage antibody library. *European Journal of Biochemistry*. 260 (3), 774–784.

Hsing, A. W. and Chokkalingam, A. P. (2006). Prostate cancer epidemiology. *Frontiers in Bioscience : A Journal and Virtual Library*, 11, 1388–1413.

Hu, X. T., Chen, W., Wang, D., Shi, Q. L., Zhang, F. B., Liao, Y. Q., Jin, M. and He, C. (2008). The proteasome subunit PSMA7 located on the 20q13 amplicon is overexpressed and associated with liver metastasis in colorectal cancer. *Oncology Reports*, 19 (2), 441–446.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4 (1), 44–57.

Humphrey, P. A. (2004). Gleason grading and prognostic factors in carcinoma of the prostate. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc*, 17 (3), 292–306.

Irish Cancer Society (2015). *Irish Cancer Society* [online]. Available from: http://www.cancer.ie/ [Accessed 27 March 2015].

Irish Cancer Society (2013). *Understanding Early Prostate Cancer* [online]. Available from: http://www.cancer.ie/sites/default/files/content-attachments/understanding_early_prostate_cancer.pdf [Accessed 30 March 2015].

Jacobs, E. J., Rodriguez, C., Mondul, A. M., Connell, C. J., Henley, S. J., Calle, E. E. and Thun, M. J. (2005). A large cohort study of aspirin and other nonsteroidal anti-inflammatory drugs and prostate cancer incidence. *Journal of the National Cancer Institute*, 97 (13), 975–980.

Jalloh, M., Myers, F., Cowan, J. E., Carroll, P. R. and Cooperberg, M. R. (2014). Racial Variation in Prostate Cancer Upgrading and Upstaging Among Men with Low-risk Clinical Characteristics. *European Urology*, 67 (3), 451–457.

Janeway, C. A., Travers, P., Walport, M. and Shlomchik, M. (2001). *Immunobiology: The Immune System In Health And Disease*, 5th Edition. Garland Science, New York.

Jespersen, C. G., Nørgaard, M., Friis, S., Skriver, C. and Borre, M. (2014). Statin use and risk of prostate cancer: A Danish population-based case-control study, 1997-2010. *Cancer Epidemiology*, 38 (1), 42–47.

Jones, M. L., Alfaleh, M. A., Kumble, S., Zhang, S., Osborne, G. W., Yeh, M., Arora, N., Hou, J. J. C., Howard, C. B., Chin, D. Y. and Mahler, S. M. (2016). Targeting membrane proteins for antibody discovery using phage display. *Scientific Reports*, 6: 26240, 1-11.

Kantoff, P., Higano, C., Shore, N., Berger, E., Small, E., Penson, D., Redfern, C. H, Ferrari, A. C., Dreicer, R., Sims, R. B., Xu, Y., Frohlich, M. W. and Schellhammer, P. F. (2010). Sipuleucel-T Immunotherapy for Castration-Resistant Prostate Cancer. *New England Journal of Medicine,* 363, 411-422.

Karantanos, T., Evans, C. P., Tombal, B., Thompson, T. C., Montironi, R. and Isaacs, W. B. (2015). Understanding the Mechanisms of Androgen Deprivation Resistance in Prostate Cancer at the Molecular Level, *European Urology*, 67 (3), 470–479.

Keller, E. T. (2007). Biology and therapeutic basis of prostate cancer bone metastasis in: Chung, L. W. K., Isaacs, W. B., and Simons, J. W. (eds.) *Prostate Cancer.* Second edition. Totowa, NJ, USA: Humana Press, 175–191.

Keller, T., Kalt, R., Raab, I., Schachner, H., Mayrhofer, C., Kerjaschki, D. and Hantusch, B. (2015). Selection of scFv Antibody Fragments Binding to Human Blood versus Lymphatic Endothelial Surface Antigens by Direct Cell Phage Display. *PLoS ONE*, 10 (5), e0127169.

Kibat, J., Schirrmann, T., Knape, M. J., Helmsing, S., Meier, D., Hust, M., Schroder, C.,

Bertinetti, D., Winter, G., Pardes, K., Funk, M., Vala, A., Giesse, N., Herberg, F. W., Dubel, S. and Hoheisel, J. D. (2015). Utilisation of antibody microarrays for the selection of specific and informative antibodies from recombinant library binders of unknown quality. *New Biotechnology*, 33 (5), 574-581.

Kijanka, G., Barry, R., Chen, H., Gould, E., Seidlits, S. K., Schmid, J., Morgan, M., Mason, D. Y., Cordell, J. and Murphy, D. (2009a). Defining the molecular target of an antibody derived from nuclear extract of Jurkat cells using protein arrays. *Analytical Biochemistry*, 395 (2), 119–124.

Kijanka, G., Hector, S., Kay, E. W., Murray, F., Cummins, R., Murphy, D., MacCraith, B. D., Prehn, J. H. M. and Kenny, D. (2010). Human IgG antibody profiles differentiate between symptomatic patients with and without colorectal cancer, *Gut*, 59 (1), 69–78.

Kijanka, G., IpCho, S., Baars, S., Chen, H., Hadley, K., Beveridge, A., Gould, E. and Murphy, D. (2009b). Rapid characterization of binding specificity and cross-reactivity of antibodies using recombinant human protein arrays. *Journal of Immunological Methods*, 340 (2), 132–137.

Kijanka, G. and Murphy, D. (2009). Protein arrays as tools for serum autoantibody marker discovery in cancer. *Journal of Proteomics*, 72 (6), 936–944.

Kim, K. M., Kang, M. and Yi, E. C. (2011). Applications of cell-based phage display panning to proteomic analysis. *Genes & Genomics*, 33 (1), 9–15.

Köhler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity., *Nature*, 256 (5517), 495–497.

Krah, S., Schröter, C., Zielonka, S., Empting, M., Valldorf, B. and Kolmar, H. (2016). Single-domain antibodies for biomedical applications. *Immunopharmacology and Immunotoxicology*, 38 (1), 21–28.

Laemmli, U. (1970). Cleavage of structural protins during the assembly of the head of

bactephage T4. *The Journal of Biological Chemistry*, 227, 1–6.

Lemass, D., Hearty, S. and O'Kennedy, R. (2014). Surface Plasmon Resonance-Based Immunoassays, in: Koistinen, H. and Stenman, U.-H. (eds.) *Novel Approaches in Immunoassays*. London: Future Science Group, 20–40.

Lemass, D., O'Kennedy, R. and Kijanka, G. S. (2016). Referencing cross-reactivity of detection antibodies for protein array experiments [version 1; referees; 1 approved, 2 approved with reservations]. *F1000Research*, 5:73.

Leonard, P., Säfsten, P., Hearty, S., McDonnell, B., Finlay, W. and O'Kennedy, R. (2007). High throughput ranking of recombinant avian scFv antibody fragments from crude lysates using the Biacore A100. *Journal of Immunological Methods*, 323 (2), 172–179.

Lilja, H., Ulmert, D. and Vickers, A. J. (2008). Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nature Reviews. Cancer*, 8 (4), 268–278.

Lim, J. P. and Gleeson, P. A. (2011). Macropinocytosis: an endocytic pathway for internalising large gulps. *Immunology and Cell Biology*, 89 (8), 836–843.

Lin, H.-K., Altuwaijri, S., Lin, W.-J., Kan, P.-Y., Collins, L. L. and Chang, C. (2002). Proteasome activity is required for androgen receptor transcriptional activity via regulation of androgen receptor nuclear translocation and interaction with coregulators in prostate cancer cells. *The Journal of Biological Chemistry*, 277 (39), 36570–36576.

Little, M., Kipriyanov, S. M., Le Gall, F. and Moldenhauer, G. (2000). Of mice and men: hybridoma and recombinant antibodies., *Immunology Today*, 21 (8), 364–370.

Liu, J. K. H. (2014). The history of monoclonal antibody development - Progress, remaining challenges and future innovations. *Annals of Medicine and Surgery,* 3 (4), 113-116.

Lu, M., Faull, K. F., Whitelegge, J. P., He, J., Shen, D., Saxton, R. E. and Chang, H. R. (2007).

Proteomics and mass spectrometry for cancer biomarker discovery. *Biomarker Insights*, 2, 347–360.

Mabjeesh, N. J., Willard, M. T., Frederickson, C. E., Zhong, H. and Simons, J. W. (2003). Androgens stimulate hypoxia-inducible factor 1 activation via autocrine loop of tyrosine kinase receptor/phosphatidylinositol 3'-kinase/protein kinase B in prostate cancer cells. *Clinical Cancer Research*, 9 (7), 2416–2425.

Maglott, D., Ostell, J., Pruitt, K. D. and Tatusova, T. (2011). Entrez Gene: gene-centered information at NCBI., *Nucleic Acids Research*, 39, D52-7.

Martin, A. C. R. (2015). *Antibodies* [online]. Available from: http://www.bioinf.org.uk/abs/#cdrid [Accessed 16 March 2015].

Mayor, S., Parton, R. G. and Donaldson, J. G. (2014). Clathrin-independent pathways of endocytosis. *Cold Spring Harbor Perspectives in Biology*, 6 (6), a016758.

McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, 348 (6301), pp. 552–554.

Mottet, N., Bellmunt, J., Briers, E., Bergh, R. C. N. van den, Bolla, M., van Casteren, N. J., Conford, P., Culine, S., Joniau, S., Lam, T., Mason, M. D., Matveev, V., van der Poel, H., van der Kwast, T. H., Rouviere, O. and Wiegel, T. (2015). Guidelines on Prostate Cancer. *European Association of Urology,* 1-137.

Murphy, M. A., O'Connell, D. J., O'Kane, S. L., O'Brien, J. K., O'Toole, S., Martin, C., Sheils, O., O'Leary, J. J. and Cahill, D. J. (2012). Epitope presentation is an important determinant of the utility of antigens identified from protein arrays in the development of autoantibody diagnostic assays, *Journal of Proteomics*, 75 (15), 4668–4675.

National Cancer Institute (2014). Stages of Prostate Cancer, [online]. Available from: http://www.cancer.gov/cancertopics/pdq/treatment/prostate/Patient/page2

[Accessed 31 March 2015].

National Cancer Institute (2015). Prostate Cancer Treatment - Stage Information for Prostate Cancer [online]. Available from: http://www.cancer.gov/cancertopics/pdq/treatment/prostate/HealthProfessional/page2 [Accessed 24 April 2015].

National Center for Biotechnology Information (2012). How does the prostate work? [online]. Available from: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0072475/ [Accessed 1 March 2015].

Nelson, W. G., Marzo, A. M. De and Isaacs, W. B. (2003). Prostate cancer. *Lancet*, 349, 366–381.

Nemeth, J. A., Harb, J. F., Barroso, U., He, Z., Grignon, D. J. and Cher, M. L. (1999). Severe combined immunodeficient-hu model of human prostate cancer metastasis to human bone. *Cancer Research*, 59 (8), 1987–1993.

Okotie, O. T., Roehl, K. A., Han, M., Loeb, S., Gashti, S. N. and Catalona, W. J. (2007). Characteristics of prostate cancer detected by digital rectal examination only. *Urology*, 70 (6), 1117–1120.

Omidfar, K. and Daneshpour, M. (2015). Advances in phage display technology for drug discovery. *Expert Opinion on Drug Discovery*, 10 (6), 651–669.

O'Reilly, J. A., Fitzgerald, J., Fitzgerald, S., Kenny, D., Kay, E. W., O'Kennedy, R. and Kijanka, G. S. (2015). Diagnostic potential of zinc finger protein-specific autoantibodies and associated linear B-cell epitopes in colorectal cancer. *PLoS ONE*, 10 (4), 1–13.

Paner, G. P., Luthringer, D. J. and Amin, M. B. (2008). Best practice in diagnostic immunohistochemistry: prostate carcinoma and its mimics in needle core biopsies. *Archives of Pathology & Laboratory Medicine*, 132 (9), 1388–1396.

Parker, C., Nilsson, S., Heinrich, D., Helle, S. I., O'Sullivan, J. M., Fosså, S. D., *et al.* (2013). Alpha emitter radium-223 and survival in metastatic prostate cancer., *The New England Journal of Medicine*, 369 (3), 213–23.

Pavelka, M. and Roth, J. (2010). Fluid-Phase Endocytosis and Phagocytosis, in: *Functional Ultrastructure*. Vienna: Springer Vienna, 104–105.

Pearson, W. R. (2013). Selecting the Right Similarity-Scoring Matrix. *Current Protocols in Bioinformatics*, 43, 3.5.1-3.5.9.

Persson, H., Ye, W., Wernimont, A., Adams, J. J., Koide, A., Koide, S., Lam, R. and Sidhu, S. S. (2013). CDR-H3 Diversity Is Not Required for Antigen Recognition by Synthetic Antibodies. *Journal of Molecular Biology*, 425 (4), 803–811.

Phin, S., Moore, M. W. and Cotter, P. D. (2013) Genomic Rearrangements of PTEN in Prostate Cancer. *Frontiers in Oncology*, 3 (September), 240.

Popkov, M., Rader, C. and Barbas, C. F. (2004). Isolation of human prostate cancer cell reactive antibodies using phage display technology. *Journal of Immunological Methods*, 291, 137–151.

Qiagen (2003). *A handbook for high-level expression and purification of 6xHis-tagged proteins*, *The QIAexpressionist* [online] Available from: https://www.qiagen.com/mx/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en [Accessed 5 December 2015].

Rasband, W.S (2012). ImageJ, *U. S. National Institutes of Health, Bethesda, Maryland, USA* [online] Availible from: http://imagej.nih.gov/ij/ [Accessed 10 May 2016].

Ratcliffe, M. J. (2008). *B Cells, the Bursa of Fabricius and the Generation of Antibody Repetoires*, Davison, F., Kaspers, B., and Schat, K. A. (eds.) *Avian Immunology*. First ed. Amsterdam, The Netherlands, 101-118.

Reichert, J. M. (2015). Therapeutic monoclonal antibodies approved or in review in the European Union or United States [online]. Available from: http://www.antibodysociety.org/news/approved_mabs.php [Accessed 23 April 2015].

Richter, F., Joyce, A., Fromowitz, F., Wang, S., Watson, J., Watson, R., Irwin Jr., R. J. and Huang, H. F. (2002). Immunohistochemical localization of the retinoic Acid receptors in human prostate. *Journal of Andrology*, 23 (6), 830–838.

Rudnick, S. I., Lou, J., Shaller, C. C., Tang, Y., Klein-Szanto, A. J. P., Weiner, L. M., Marks, J. D. and Adams, G. P. (2011). Influence of affinity and antigen internalization on the uptake and penetration of Anti-HER2 antibodies in solid tumors. *Cancer Research*, 71 (6), 2250–2259.

Saad, F., Gleason, D. M., Murray, R., Tchekmedyian, S., Venner, P., Lacombe, L., Chin, J. L., Goas, J.A. and Zheng, M. (2004). Long-term efficacy of zoledronic acid for the prevention of skeletal complications in patients with metastatic hormone-refractory prostate cancer. *Journal of the National Cancer Institute*, 96 (11), 879–882.

Schägger, H. and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166 (2), 368–379.

Scher, H. I., Cabot, R. C., Harris, N. L., Rosenberg, E. S., Shepard, J.-A. O., Cort, A. M., *et al.* (2012). Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. *New England Journal of Medicine*, 367 (13), 1187–1197.

Schroeder, H. W., Cavacini, L. and Cavacini, L. (2010). Structure and function of immunoglobulins. *The Journal of Allergy and Clinical Immunology*, 125 (2), S41-52.

Schulman, C. C. and Ekane, S. (2001). Nutrition and Prostate Cancer. *Urology*, 4295 (1), 318–334.

Shah, R. B. (2009). Current perspectives on the Gleason grading of prostate cancer.

Archives of Pathology & Laboratory Medicine, 133 (11), 1810–1816.

Sharma, P., Hu-Lieskovan, S., Wargo, J. A. and Ribas, A. (2017). Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell*, 168 (4), 707–723.

Sharma, S., Zapatero-Rodríguez, J., Estrela, P. and O'Kennedy, R. (2015). Point-of-Care Diagnostics in Low Resource Settings: Present Status and Future Role of Microfluidics. *Biosensors*, 5 (3), 577–601.

Shen, Z., Stryker, G. A., Mernaugh, R. L., Yu, L., Yan, H. and Zeng, X. (2005). Single-chain fragment variable antibody piezoimmunosensors. *Analytical Chemistry*, 77 (3), 797–805.

Shih, H. H., Tu, C., Cao, W., Klein, A., Ramsey, R., Fennell, B. J., Lambert, M., Ni Shuilleabhain, D., Autin, B., Kouranova, E., Laxmanan, S., Braithwaite, S., Wu, L., Ait-Zahra, M., Dumin, A. J., LaVallie, E. R., Arai, M., Corcoran, C., Paulsen, J. E., Gill, D., Cunningham, O., Bard, J., Mosyak, L. and Finlay, W. J. (2012). An ultra-specific avian antibody to phosphorylated tau protein reveals a unique mechanism for phosphoepitope recognition. *Journal of Biological Chemistry*, 287 (53), 44425–44434.

Shukla, G. S. and Krag, D. N. (2010). Cancer cell-specific internalizing ligands from phage displayed β-lactamase-peptide fusion libraries. *Protein Engineering, Design and Selection*, 23 (6), 431–440.

Siegel, R. L., Miller, K. D. and Jemal, A. (2015). Cancer Statistics, 2015. *CA: A Cancer Journal for Clinicians*, 65 (1), 5–29.

Siva, A. C., Kirkland, R. E., Lin, B., Maruyama, T., McWhirter, J., Yantiri-Wernimont, F., Bowdish, K. S. and Xin, H. (2008). Selection of anti-cancer antibodies from combinatorial libraries by whole-cell panning and stringent subtraction with human blood cells. *Journal of Immunological Methods*, 330, 109–119.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface., *Science (New York, N.Y.)*, 228 (4705), 1315–1317.

Source Bioscience (2010). *Protein Macroarrays Manual* [online]. Available from: http://www.lifesciences.sourcebioscience.com/media/290406/sbs_ig_manual_protein array_v1.pdf [Accessed 11 November 2015]

Steeland, S., Vandenbroucke, R. E. and Libert, C. (2016). Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discovery Today*, 21 (7), 1076–1113.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41 (1), 207–234.

Swallow, T., Chowdhury, S. and Kirby, R. S. (2012). Cancer of the prostate gland. *Medicine*, 40 (1), 10–13.

Tate, J. and Ward, G. (2004). Interferences in Immunoassay. *The Clinical Biochemist Reviews*, 25 (2), 105-120.

Thalmann, G. N., Anezinis, P. E., Chang, S., Thaimann, N., Hopwood, V. L., Pathak, S., Eschenbach, A. Von and Chung, L. K. (1994). Androgen-independent Cancer Progression and Bone Metastasis in the LNCaP Model of Human Prostate Cancer. *Cancer Research*, 54, 2577–2581.

Townsend, S., Fennell, B. J., Apgar, J. R., Lambert, M., McDonnell, B., Grant, J., Wade, J. Franklin, J., Foy, N., Ni Shuilleabhain, D., Fields, C., Darmanin-Sheehan, A., King, A., Paulsen, J. E., Hickling, T. P., Tchitiakova, L., Cunningham, O. and Finlay, W. J. (2015). Augmented Binary Substitution: Single-pass CDR germ-lining and stabilization of therapeutic antibodies. *Proceedings of the National Academy of Sciences of the United States of America*, 112 (50), 15354–15359.

Uhlen, M., Bandrowski, A., Carr, S., Edwards, A., Ellenberg, J., Lundberg, E., Rimm D. L., Rodriguez, H., Hiltke, T., Synder, M. and Yamamoto, T. (2016). A proposal for validation of antibodies, *Nature Methods*, 13 (10), 823-827.

Wang, L.-N., Tong, S.-W., Hu, H.-D., Ye, F., Li, S.-L., Ren, H., Zhang, D.-Z., Xiang, R. and Yang, Y.-X. (2012). Quantitative proteome analysis of ovarian cancer tissues using a iTRAQ approach. *Journal of Cellular Biochemistry*, 113 (12), 3762–3772.

Wang, X., Yu, J., Sreekumar, A., Varambally, S., Shen, R., Giacherio, D., Mehra, R., Montie, J. E., Pienta, K. J., Sanda, M. G., Kantoff, P. W., Rubin, M. A., Wei, J. T., Ghosh, D. and Chinnaiyan, A. M. (2005) Autoantibody signatures in prostate cancer., *The New England Journal of Medicine*, 353 (12), 1224–1235.

Warr, G. W., Magor, K. E. and Higgins, D. A. (1995). IgY: clues to the origins of modern antibodies. *Immunology Today*, 16 (8), 392–398.

Weinstein Dunn, M. and Wallace Kazer, M. (2011). Prostate cancer overview. *Seminars in Oncology Nursing*, 27 (4), 241–250.

Woof, J. M. and Burton, D. R. (2004). Human antibody-Fc receptor interactions illuminated by crystal structures. *Nature Reviews Immunology*, 4, 89–99.

Wride, M. A., Mansergh, F. C., Somani, J. M., Winkfein, R. J. and Rancourt, D. E. (2002). Characterization and *in silico* mapping of a novel murine zinc finger transcription factor. *Gene*, 289 (1–2), 49–59.

Wu, L., Oficjalska, K., Lambert, M., Fennell, B. J., Darmanin-Sheehan, A., Ní Shúilleabháin, D., Autin, B., Cummins, E., Tchisiakova, L., Bloom, L., Gill, D., Cunningham, O. and Finlay, W. J. (2012). Fundamental Characteristics of the Immunoglobulin VH Repertoire of Chickens in Comparison with Those of Humans, Mice, and Camelids. *The Journal of Immunology*, 188 (1), 322–333.

Xu, J. L. and Davis, M. M. (2000). Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity*, 13 (1), 37–45.

Yamaguchi, H. and Condeelis, J. (2007). Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochimica et Biophysica Acta (BBA) - Molecular Cell*

Research, 1773 (5), 642–652.

Zhang, X. and Oglesbee, M. (2003). Use of surface plasmon resonance for the measurement of low affinity binding interactions between HSP72 and measles virus nucleocapsid protein. *Biological Procedures Online*, 5, 170–181.

Zhou, Y., Zhao, L. and Marks, J. D. (2012). Selection and characterization of cell binding and internalizing phage antibodies. *Archives of Biochemistry and Biophysics*, 526 (2), 107–113.

Chapter 8 - Appendix

Appendix A – Protein array raw images

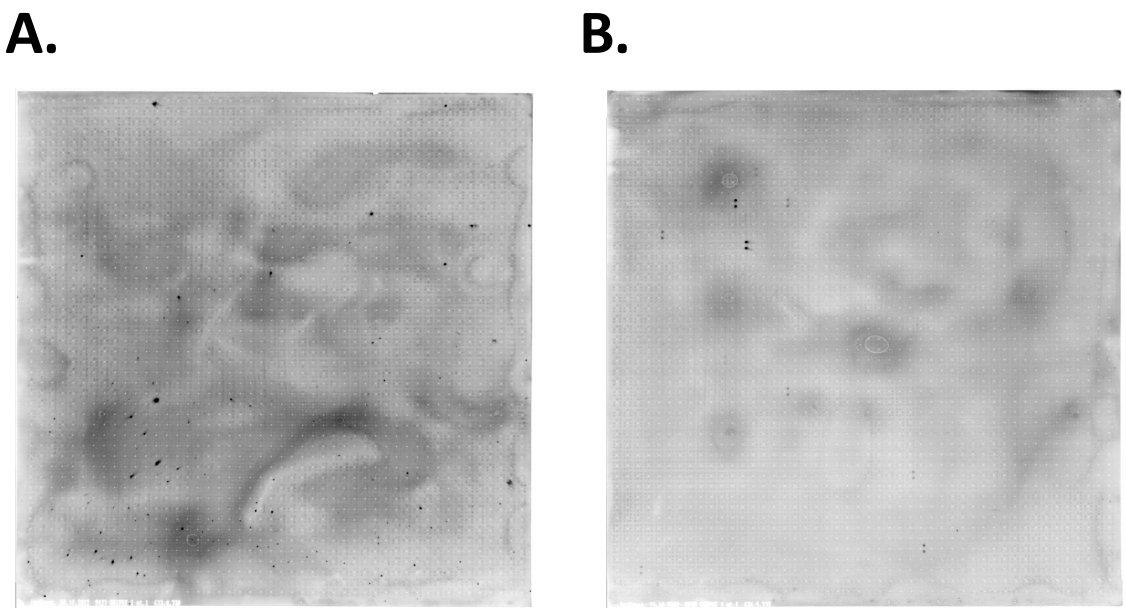

Figure 8.1 - Raw image data obtained from array profiling of avian serum secondary antibodies. See Figure 4.2 for details.

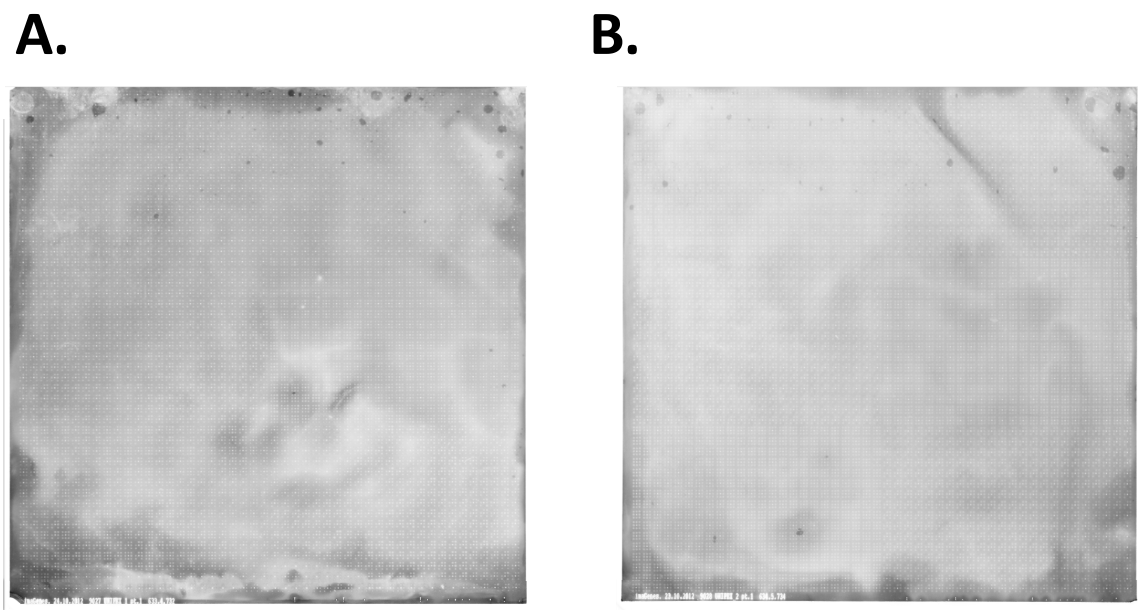

Figure 8.2 - Raw image data obtained from array profiling of avian pre-immunisation bleed serum prior to a PCa cell immunisation regime See Figure 4.5 for details.

A. B.

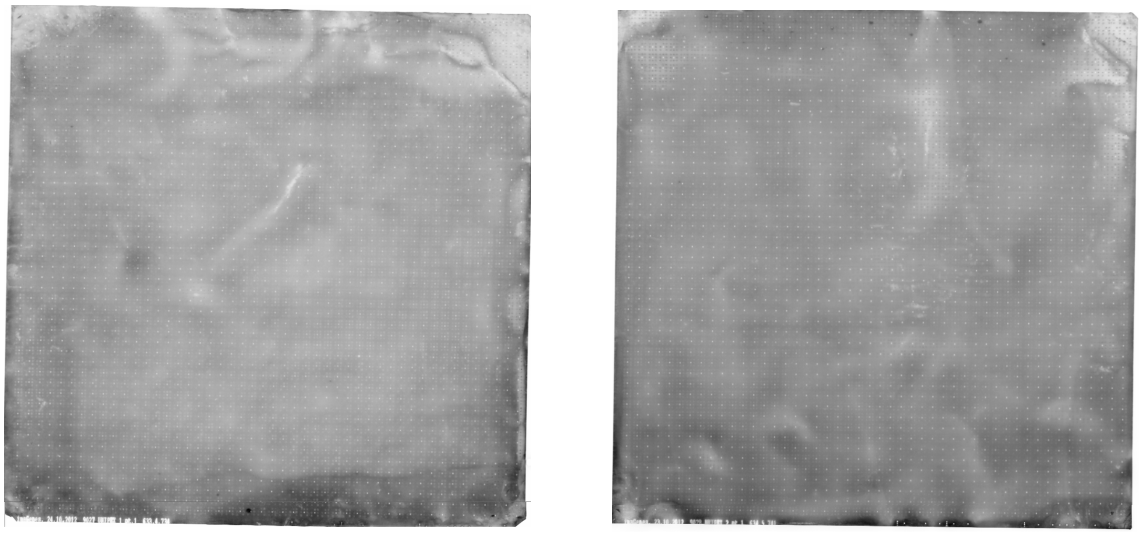

Figure 8.3 - Raw image data obtained from array profiling of avian post-immunisation serum following a PCa cell immunisation regime See Figure 4.6 for details.

A. B.

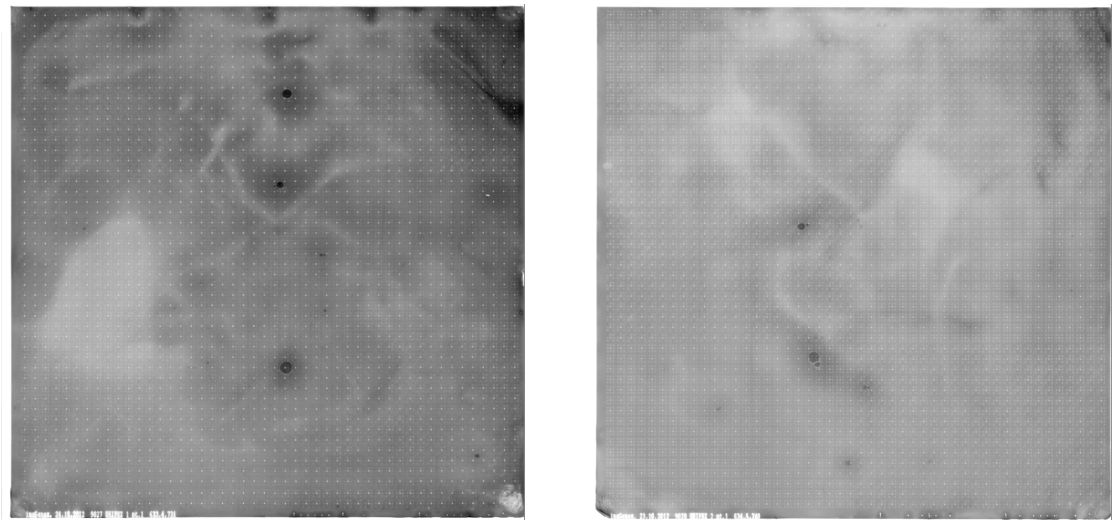

Figure 8.4 - Raw image data obtained from array profiling of anti-bacteriophage secondary antibodies. See Figure 4.3 for details.

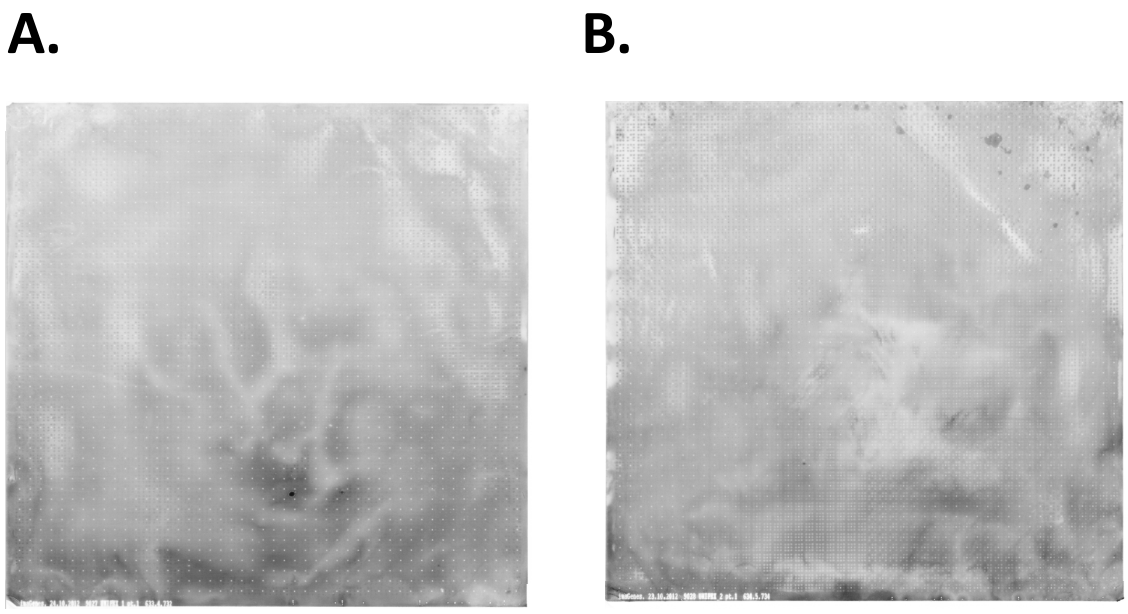

Figure 8.5 - Raw images showing protein array antigen profiling of an avian scFv-phage library generated against PCa cells. See Figure 4.7 for details.

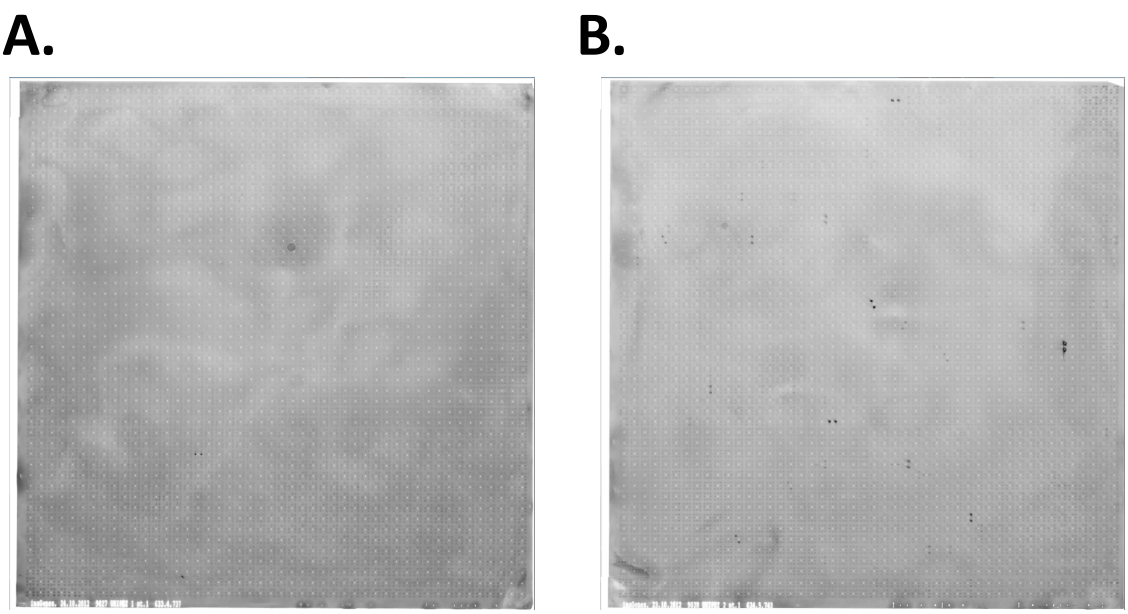

Figure 8.6 – Raw image data obtained from array profiling of scFv secondary antibodies. See Figure 4.4 for details.

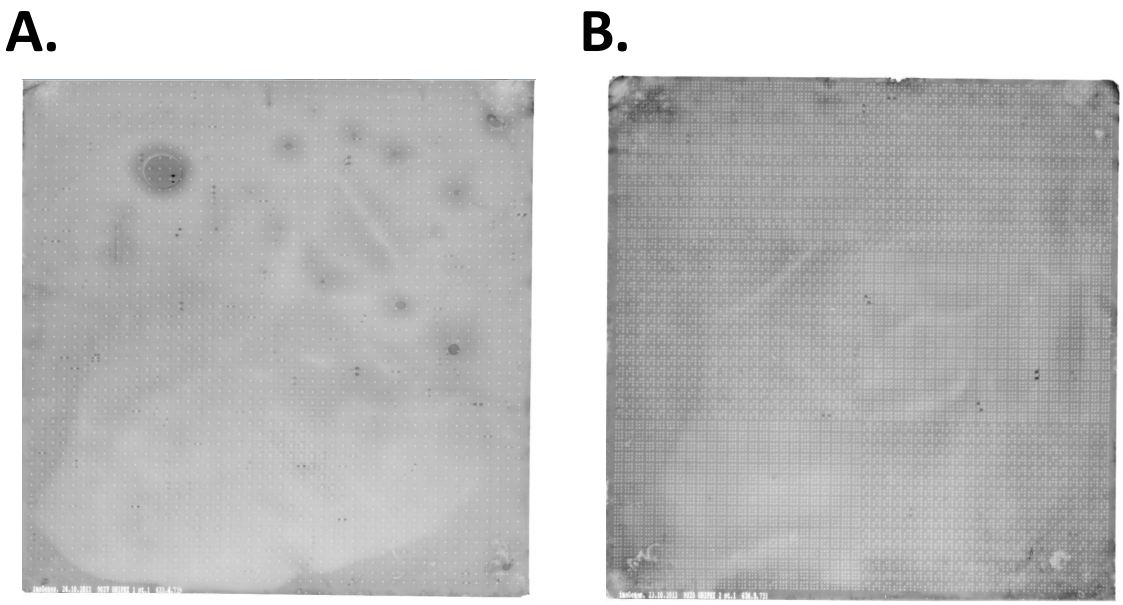

Figure 8.7 - Raw images showing protein array profiling of scFv from 384 clones enriched against PCa cells via cell panning. See Figure 4.8 for details

Appendix B – BLAST and SIM analysis

Table 8.1 –BLAST and SIM analysis of rabbit anti-IgY and AP-conjugated goat anti-rabbit false positives compared against four chicken and four rabbit immunoglobulin proteins for homology. Included in the table is each antigen's similarity by BLAST E-value and SIM score. 'NSSF' refers to 'No significant similarity found'. See section 4.4.2 for further details.

Gen Bank ID	Array Signal Intensity	Chicken IgY Gene comparison								Rabbit IgG Gene comparison							
		Light Chain				Heavy Chain				Light Chain				Heavy Chain			
		Ig lambda chain C region		Ig lambda chain V-1 region		immunoglobulin Y heavy chain constant region		immunoglobulin Y heavy chain variable region		Ig lambda chain C region		Ig lambda chain variable region, partial		Ig gamma chain C region		immunoglobulin heavy chain VDJ region, partial	
		BLAST E - value	SIM high score value	BLAST E - value	SIM high score value	BLAST E – value	SIM high score value	BLAST E - value	SIM high score value	BLAST E - value	SIM high score value	BLAST E - value	SIM high score value	BLAST E - value	SIM high score value	BLAST E - value	SIM high score value
BM914329	3	6.00E-04	42	2.00E-04	53	0.4	43	1.00E-48	393	3.00E-05	58	3.00E-06	76	3.00E-10	114	3.00E-47	356
BQ711793	3	NSSF	23	3.00E-06	50	0.25	36	4.00E-30	210	NSSF	23	NSSF	39	0.59	26	1.00E-29	194
BQ709082	3	6.2	22	4.00E-10	48	2.7	32	2.00E-34	218		29	2.00E-05	52	0.27	36	1.00E-41	291
BX417981	3	3.00E-13	86	5.5	24	1.3	34	0.027	36	5.00E-17	137	NSSF	34	2.00E-125	890	0.044	36
118471	3	3.6	24	NSSF	21	NSSF	29	0.094	25	1	26	NSSF	25	2.6	28	0.48	22
BC044933	2	9.1	30	0.9	31	1.2	37	0.27	26	1.1	32	NSSF	29	0.13	41	3.8	27

BC010132	2	3.3	33	1.7	27	3.1	34	0.5	31	0.29	25	1.7	34	1.2	36		27
AK092483	2	0.11	27	0.86	32	0.84	34	0.71	36	0.66	30	NSSF	24	5.6	26	NSSF	23
NM_006814	2	NSSF	29	9.2	29	0.26	33	NSSF	24	1.3	27	7.6	24	0.48	31	4.5	23
DA970556	2	0.71	23	1.00E-06	39	0.2	45	3.00E-44	313	0.16	27	4.00E-08	64	0.009	41	3.00E-49	357
NM_002228	2	NSSF	22	2.00E-04	39	0.42	33	NSSF	31	NSSF	24	NSSF	30	0.009	49	NSSF	32
BC041022	2	NSSF	27	NSSF	26	NSSF	38	NSSF	24	9.5	24	NSSF	30	1.5	36	NSSF	21
BC018708	2	NSSF	26	NSSF	24	0.24	44	1.8	37	NSSF	25	1.1	28	1.3	32	0.32	27
BC018708	2	NSSF	26	NSSF	26	0.25	44	NSSF	37	NSSF	25	4.3	28	1.8	36	1.3	27
BM920476	2	0.24	31	0.16	49	0.12	51	NSSF	26	5.5	30	6	39	0.051	49	6	39
BX417981	2	NSSF	28	2.00E-04	54	2.1	36	6.00E-22	200	0.025	40	7.00E-23	213	NSSF	34	0.033	54
BG754662	2	4.00E-14	122	2.3	33	5.7	32	1.9	30	4.00E-148	1033	0.37	33	2.00E-16	146	NSSF	31
NM_001978	1	0.47	27	0.3	33	0.15	37	1.3	27	0.14	34	0.92	24	6.1	32	NSSF	NSSF
NM_080881	1	NSSF	35	3.8	26	1.3	35	NSSF	29	4.2	32	4.8	31	0.38	34	NSSF	NSSF

NM_001012426	1	2.5	32	4.3	30	1.7	39	3.4	31	0.88	31	5.5	33	1.6	32	0.95	26
BX647115	1	0.74	30	0.13	36	2	31	4.7	29	2.1	28	1	27	5.4	33	0.27	35
AF479827	1	7.9	28	4.1	31	0.36	48	0.7	37	1.9	26	0.9	30	0.6	40	NSSF	28
NM_022489	1	1.1	30	NSSF	28	3.6	41	8.1	36	0.39	30	0.15	38	1.2	51	3.4	28
BC000786	1	0.37	34	NSSF	48	NSSF	28	NSSF	27	1.3	27	NSSF	26	NSSF	31	1.5	29
BC008343	1	NSSF	31	0.8	33	0.62	36	0.16	34	0.053	37	NSSF	31	0.89	34	1.5	28
BC000459	1	NSSF	25	0.23	33	0.91	34	NSSF	35	3.6	26	NSSF	30	0.24	36	0.9	27
BC040880	1	NSSF	31	NSSF	25	0.24	51	0.85	29	NSSF	23	NSSF	29	0.61	27	0.38	27
NM_003260	1	3.9	28	NSSF	29	1.6	41	NSSF	28	0.064	33	0.48	30	0.24	32	0.86	31
AL833379	1	0.56	32	NSSF	22	3.7	33	1.5	27	4.8	22	NSSF	23	4.3	26	NSSF	29
NM_006548	1	NSSF	26	1.7	32	NSSF	30	2.8	26	NSSF	30	1.8	25	0.25	42	NSSF	28
AK097073	1	0.001	39	8.1	28	1	30	1.5	29	NSSF	27	0.17	32	NSSF	32	NSSF	24
AK127401	1	1.3	29	NSSF	28	NSSF	32	0.51	38	2	24	2.5	28	0.71	30	1.5	32

BC036307	1	NSSF	30	1.5	31	3.4	36	0.56	27	5.8	28	NSSF	28		31	0.22	29
BC146654	1	1.4	37	NSSF	30	0.035	52	3.6	42	1.1	32	1.8	33	0.99	35	3.1	31
BC022890	1	0.013	36	0.1	31	0.072	41	5.8	26	0.62	31	NSSF	27	0.58	35	2.3	23
NM_002087	1	2	26	5.5	27	2.5	28	2.2	27	NSSF	26	3.5	26	8.5	36	1.4	27
NM_032627	1	5.8	23	4.4	33	5.9	56	0.45	38	NSSF	27	5.9	26	0.44	40	0.6	32
AK127255	1	1.3	33	NSSF	31	0.73	54	1.3	31	4.3	26	0.47	34	0.82	39	1.2	34
BC090883	1	NSSF	27	NSSF	25	NSSF	61	0.38	27	NSSF	24	NSSF	26	NSSF	32	1.2	26
NM_003768	1	0.1	31	0.97	24	0.095	35	NSSF	23	0.23	28	NSSF	23	0.045	35	3.7	25
AB002328	1	0.29	37	0.25	40	0.9	41	4.1	30	0.18	40	0.17	41	0.088	43	NSSF	32
AK096320	1	1.5	27	NSSF	25	NSSF	29	NSSF	26	0.062	33	4.9	27	0.27	35	NSSF	33
NM_015695	1	4.8	26	0.38	36	0.39	41	2.9	31	3.6	33	0.8	32	1.6	38	NSSF	24
NM_007371	1	NSSF	30	2.7	46	NSSF	35	NSSF	29	NSSF	34	2.6	34	6.5	41	0.75	29
NM_014866	1	0.73	27	0.28	35	0.3	45	4.7	24	2.4	26	0.28	32	1.3	35	4.9	37

NM_031372	1	7.4	23	0.08	30	2.4	35	2.9	32	NSSF	23	NSSF	28	3.1	29	2.2	29
NM_014811	1	0.16	30	7.7	33	1.4	59	0.053	45	0.3	31		34	2.9	35	0.78	33
NM_001098800	1	4.6	27	5.2	40	NSSF	33	0.24	34	1.7	32	0.041	41	2.5	34	9.2	25
BC037307	1	3.9	33	2.2	31	7.9	41	0.85	38	3	34	4.5	35	2.8	40	2.7	44
AB208876	1	0.78	34	4.7	32	0.065	46	0.77	31	NSSF	26	1.7	29	3	35	0.67	31
BF110897	1	Sequence Unavailable															
BC004352	1	0.71	30		24	0.087	46		23		25	0.071	38	2.1	33	NSSF	25
AK225632	1	0.93	32	3.6	30	0.24	43	8.4	36	NSSF	32	3.6	37	1.2	31	8.9	30
AL133055	1	5.5	36	3.8	26	NSSF	44		38	0.96	30	0.64	38		32	2.6	29
NM_032329	1		24		30	NSSF	31	2.9	23	NSSF	23		23	5.2	31	2.9	32
XR_015693	1	Sequence Unavailable															
BC006105	1	1.8	30	NSSF	26	2.3	50	1.8	28	NSSF	25	NSSF	32	0.22	34	8.1	23
NM_002140	1	3.4	30	0.13	34	4	33	0.11	33	2.4	32	1.8	36	4.6	37	0.76	31
NM_001270	1	6	32	0.17	41	7	49	1.7	31	7.7	32	4.5	38	1.5	34	0.87	33

AK124880	1	0.092	37	NSSF	27	0.46	49	2	25	1.6	29	1.2	34	2.4	36	1.7	25
AB209272	1	1.1	27	1	28		31	4.6	32		23	4.3	26	8.9	29	2.4	25
CR606311	1	Sequence Unavailable															
AK128584	1	NSSF	29	0.35	35	NSSF	38	NSSF	30	NSSF	27	5.7	32	NSSF	33	1.8	30

Appendix C – Protein array-identified antigens

Table 8.2 – GenBank IDs of false positive signals from control experiments. Detection antibodies used in sections 4.2.2, 4.2.3 and 4.2.4 were profiled in the absence of their manufacturer-listed antigen. Included in the table is a 'Shared' column, which includes overlapping false positives that can be attributed to the AP-conjugated goat anti-rabbit antibody. GenBank IDs highlighted in bold font were present in and removed from the final experimental lists as false positives.

Serum		Phage	scFv		Shared
BM914329	BC036307	AK225761	**BQ940058**	NM_001085454	AL833379
BQ711793	BC146654	NM_017453	**BC048198**	BC017045	NM_002078
BQ709082	BC022890	BC008343	**NM_017453**	AK097873	BC008343
BX417981	NM_002087	NM_014321	**BC048198**	BC142941	BC092520
BC044933	NM_032627	BC092520	**AK226177**	BX647981	NM_014321
BC010132	AK127255	NM_002078	AB041012	NM_020780	NM_017453
AK092483	BC090883	NM_014615	NM_144563	NM_005506	NM_014615
NM_006814	NM_003768		NM_000037	NM_024297	
DA970556	AB002328		BQ051441	NM_014321	
NM_002228	AK096320		NM_152345	BX647569	
BC041022	NM_015695		AB235153	AL832819	
BC018708	NM_007371		BC022880	XM_496826	
BC018708	NM_014866		AK094894	AL833379	
BM920476	NM_031372		NM_002814	CR602296	
BX417981	NM_014811		NM_000983	NM_183040	
BG754662	NM_001098800		NM_203505	BC040312	
NM_001978	BC037307		NM_016194	AB208902	
NM_080881	AB208876		AK001674	NM_031482	
NM_001012426	BF110897		CR590682	XM_291277	
BX647115	BC004352		BC012347	NM_005426	
AF479827	AK225632		NM_012465	AK128553	
NM_022489	AL133055		NM_022356	BC092520	
BC000786	NM_032329		BQ279187	NM_000302	
BC008343	XR_015693		CR626583	NM_138361	
BC000459	BC006105		AK126779	NM_002078	
BC040880	NM_002140		NM_004715	BX640852	
NM_003260	NM_001270		XR_019439	BU565548	
AL833379	AK124880		BC008343	AK128154	
NM_006548	AB209272		NM_018683	AK024450	
AK097073	CR606311		AL833518	AL136829	
AK127401	AK128584		AK057117	BE791300	
			NM_014615	BC050412	

Table 8.3 – Antigen profile identified of chicken ABG055's serum IgY prior to PCa cell line immunisations using Unipex Protein Array Technology.

Array	Intensity	GenBank ID
1	1	AB208910
1	1	AI652785
1	1	BC078175
1	1	BM805249
1	1	NM_001048205
1	1	NM_001420
1	1	BC035030

Table 8.4 – Antigen profile identified from chicken ABG055's serum-contained IgY after PCa cell line immunisations using Unipex Protein Array Technology.

Unipex Array	Signal Intensity	GenBank ID
1	1	BX648030
1	1	NM_004521
1	1	BM541948
1	2	NM_005327
1	1	NM_001961
1	1	IMGSp9027G0313D
1	2	AK091578
1	1	NM_003073
1	1	AF330045
1	3	AB208910
1	2	NM_006197
1	1	BU944406
1	2	AI652785
1	1	NM_014153
1	2	BC071987
1	1	CR936770
1	2	CR592759
1	2	BF102725
1	2	NM_004287
1	1	BM546656
1	3	BC078175
1	2	BQ709225
1	2	AB209197
1	2	BC000375
1	2	AB029038

Unipex Array	Signal Intensity	GenBank ID
1	2	NM_052916
1	1	NM_001005751
1	1	CK905869
1	2	NM_001759
1	3	NM_001048205
1	2	AF388384
1	1	NM_002285
1	1	CR591317
1	2	DQ462572
1	2	AK126663
1	1	IMGSp9027E0853D
1	2	BX647488
1	1	BF132477
1	1	BM994108
1	2	BC030826
1	1	NM_014807
1	1	IMGSp9027C0556D
1	1	AK092888
1	1	BQ931262
1	2	NM_000038
1	1	BC111485
1	1	NM_001843
1	1	BX648190
1	1	BX537484
1	1	BI966745

1	2	AB018260
1	1	AL832897
1	1	AB018310
1	1	DR731217
1	1	AK095541
1	2	AK091125
1	1	AY359878
1	1	AK092265
1	1	NM_013975
1	1	BX248009
1	2	NM_016580
1	1	BE785179
1	2	AK131251
1	2	AK124613
1	2	BM467814
1	1	BX647066
1	1	NM_006184
1	2	NM_001746
1	1	BQ931262
1	1	BQ057394
1	1	NM_001042476
1	2	NM_144636
1	1	EL594131
1	1	BC040312
1	1	NM_007350
1	1	XR_018315
1	1	NR_003089
1	2	AK122681
1	1	AK098311
1	1	AL534323
1	1	NM_002376
1	1	NM_005151
1	1	CR936710
1	1	BM454192
1	1	BQ920362
1	1	AB209596
1	1	NM_013318
1	1	AB209197
1	1	NM_003348
1	1	NM_001077686
1	2	XR_017492

1	1	CR749220
1	1	NM_002291
1	1	AK127498
1	1	NM_018639
1	2	BX649155
1	2	NM_001420
1	1	BC035030
1	1	NM_020461
1	1	AY700118
1	1	BF240734
1	1	CR599555
1	1	BQ057394
1	1	CD173011
1	1	NM_080927
1	1	AB018307
1	1	AA621550
1	1	BX649135
1	1	CR749837
1	1	XR_017492
1	1	AB044661
1	1	BM462466
1	1	BC043646
1	1	BC001954
1	1	NM_022893
1	1	NM_001271
1	1	AL534323
1	1	AK127792
1	2	CX761459
1	1	BC038296
1	2	AL162068
2	1	XR_015892
2	1	NM_030912
2	2	AF073727
2	1	NM_145343
2	1	NM_001006617
2	3	AB208902
2	3	AF073727
2	1	NM_006842
2	1	BX641146
2	1	NM_005219
2	1	NM_018660

1	1	AB002360
1	1	NM_003367
1	1	NM_003021

2	1	AK054976
2	1	AK098120
2	1	NM_002473

Table 8.5 – Antigen profile identified from a scFv-phage library, generated from chicken ABG055's immune repertoire.

Unipex Array	Signal Intensity	GenBank ID
1	1	XR_017279
1	1	BC030707
1	1	NM_203434
1	1	AB014523
1	1	NM_016194
1	1	AF509494
2	1	BM903678
2	1	BC002970
2	1	AK097009
2	1	CR591317
2	1	AB208902
2	1	NM_020765
2	1	BC030788
2	1	AB208876
2	1	NM_182776
2	1	BU565548

Table 8.6 – Antigen profile of 384 scFv using protein arrays.

Unipex Array	Signal Intensity	GenBank ID
1	1	XR_015390
1	1	AI198821
1	3	AA442937
1	1	NM_016162
1	1	CA450485
1	1	CR625172
1	1	BM917378
1	1	BC111028
1	2	NM_024812
1	2	NM_001810
1	2	CR595461
1	2	AK127210
1	3	NM_206825
1	2	NM_017944

Unipex Array	Signal Intensity	GenBank ID
1	2	NM_001810
1	2	NM_003758
1	3	BF680544
1	2	AK090881
1	2	NM_152556
1	2	AL534323
1	3	NM_015517
1	1	AI937260
1	1	NM_015446
1	1	XR_015390
1	1	BC116445
1	1	AK093702
1	2	BC023991
1	3	NM_001123

1	1	NM_032776
1	1	NM_032776
1	1	NM_014335
1	2	BF790759
1	1	AK091125
1	2	NM_001349
1	3	NM_003021
1	3	AK123065
1	3	NM_005104
1	3	XR_018118
1	2	NM_182914
1	1	NM_016162
1	3	NM_015517
1	2	BC015664
1	2	NM_001002860
1	1	NM_001527
1	1	AB023162
1	3	NM_001417
1	1	CV806933
1	1	CR625172
1	1	BF790759
1	2	AK001740
1	2	NM_015517
1	1	BC037311
1	2	M84739
1	3	AF005482
1	2	NM_018270
1	1	NM_152562
1	2	XM_291095
1	3	CR593275
1	2	NM_001283
1	2	NM_005861
1	1	NM_015694
1	1	CR598577
1	2	NM_001283
1	2	NM_014765
1	3	NM_001018067
1	2	BI753789
1	2	BX647066
1	2	AF217190
1	2	NM_001530

1	1	NM_000182
1	3	BX649110
1	1	BC017222
1	2	NM_004368
1	3	NM_003021
1	1	NM_000182
1	1	NM_015258
1	1	AF452102
1	1	XR_016068
1	2	NM_015173
1	1	IMGSp9027D0455D
1	1	BC068549
1	1	NM_018083
1	1	AB020718
1	3	AK074969
1	2	BF680544
1	1	NM_006451
1	1	AK092888
1	1	BQ953270
1	1	AK096611
1	3	BF680544
1	2	NM_005861
1	2	BC013200
1	3	AK094811
1	3	CD049752
1	2	AK130546
1	2	AB209149
1	1	NM_133171
1	3	NM_138558
1	2	AK096611
1	1	NM_025081
1	1	NM_001096
1	2	AK091242
1	2	NM_004968
1	1	CD173011
1	1	AK126056
1	2	AL136591
1	2	XM_291095
1	2	AB018307
1	2	CR595461
1	3	NM_003021

1	2	AK096611
1	1	AK128870
1	2	BU570828
1	1	NM_003758
1	1	BU928938
1	1	BF680546
1	1	NM_006036
1	3	BC036694
1	1	BC034956
1	1	AK094006
1	1	AL137543
1	3	AK225761
1	2	BF698898
1	2	IMGSp9027C1040D
1	3	XR_019265
1	3	NM_001283
1	2	BC075794
1	2	NM_024744
1	3	BF790759
1	2	NM_018083
1	3	NM_017453
1	2	BC073841
1	2	CR622248
1	3	AB208876
1	3	BG528898
1	3	NM_001018067
1	2	BC060867
1	2	XR_015390
1	1	IMGSp9027G0627D
1	2	NM_003932
1	1	NM_001410
1	3	NM_001417
1	3	NM_018197
1	1	AB033004
1	1	AK074035
1	1	BC018708
1	1	AK126489
1	3	XR_018118
1	3	AK128439
1	3	NM_006160
1	2	NM_003932

1	1	BC048198
1	1	BC017222
1	1	NM_001042539
1	2	BC030201
1	1	NM_138423
1	1	BQ212107
1	1	BF569053
1	1	BX537968
1	1	AK122869
1	1	AK226177
1	2	NM_031372
1	1	CR627465
1	1	BC033113
1	1	IMGSp9027C0557D
1	1	M84739
1	1	AL534323
1	2	NM_031372
1	1	BC033113
1	1	BX647444
1	1	NM_006451
1	2	AK123065
1	2	AK126489
1	3	XR_018118
1	1	XR_016079
1	2	NM_019004
1	2	NM_001029882
1	1	NM_000966
1	1	AK097577
1	2	NM_015520
1	1	BF698898
1	2	NM_015695
1	3	NM_004284
1	3	NM_005104
1	3	AK123065
1	3	AF089854
1	2	NM_006047
1	1	AY894575
1	2	IMGSp9027B0619D
1	2	XR_018315
1	2	NM_015520
1	2	AK074842

1	2	NM_005861
1	1	BC082985
1	1	BX373836
1	3	BC008506
1	1	IMGSp9027D1046D
1	2	NM_052857
1	3	NM_001961
1	3	AK124755
1	3	NM_001123
1	1	IMGSp9027C0348D
1	2	AK094345
1	2	NM_031372
1	3	NM_032999
1	2	AI198821
1	1	AB051547
1	3	NM_001789
1	1	AA780281
1	2	NM_020947
1	3	NM_003926
1	3	BE795988

1	3	BC056909
1	1	AL136787
2	1	BQ073256
2	1	BC065280
2	3	NM_006421
2	1	M84739
2	2	BC068549
2	1	AF003837
2	3	AB208876
2	2	NM_020713
2	1	NM_003021
2	2	BG528898
2	3	AK074969
2	2	NM_005104
2	1	AK092882
2	2	NM_018083
2	3	NM_198057
2	1	NM_198057
2	3	AK074969
2	2	NM_198057

Table 8.7 – Array-identified antigens that are present in more than one dataset.

Present in	Total Overlaps	GenBank ID
Postbleed and scFv	9	
		BX647066
		AK092888
		AL534323
		CD173011
		XR_018315
		AB018307
		NM_003021
		AK091125
		NM_001961
scFv-Phage and Postbleed	2	
		CR591317
		AB208902
Postbleed and Prebleed	6	
		BC078175
		NM_001048205
		BC035030
		NM_001420
		AI652785
		AB208910
scFv-Phage and ScFv	1	
		AB208876

Table 8.8 – Array-identified antigens bound more than once in an experiment.

Dataset	Genbank ID	Binding events
scFv	NM_001283	3
scFv	NM_016162	2
scFv	NM_032776	2
scFv	NM_006451	2
scFv	CR595461	2
scFv	BC033113	2
scFv	NM_003932	3
scFv	BF680544	3
scFv	M84739	3
scFv	BG528898	2
scFv	NM_198057	3
scFv	NM_003021	4
scFv	NM_031372	3
scFv	BF698898	2
scFv	AL534323	2

scFv	NM_001123	2
scFv	BC068549	2
scFv	AK096611	3
scFv	AB208876	2
scFv	BC017222	2
scFv	NM_003758	2
scFv	NM_001018067	2
scFv	NM_001417	2
scFv	AK074969	3
scFv	AK126489	2
scFv	NM_000182	2
scFv	NM_005104	3
scFv	NM_005861	3
scFv	NM_001810	2
scFv	NM_018083	3
scFv	AK123065	3
scFv	NM_015520	2
scFv	BF790759	3
scFv	XR_015390	3
scFv	CR625172	2
scFv	NM_015517	3
scFv	AI198821	2
scFv	XM_291095	2
Postbleed	AB209197	2
Postbleed	AL534323	2
Postbleed	BQ931262	2
Postbleed	AF073727	2
Postbleed	BQ057394	2
Postbleed	XR_017492	2

Appendix D – DAVID and orthology analysis of protein lists

Table 8.9 –Orthology analysis of antigens identified from chicken ABG055's serum-contained IgY after PCa cell line immunisations using Unipex protein array technology.

Human Ensembl Gene ID	Associated Gene Name	Chicken Ensembl Gene ID	Chicken Associated Gene Name	Last common ancestor with Chicken	Similarity Percentage
ENSG00000105698	USF2				0
ENSG00000104915	STX10				0
ENSG00000167371	PRRT2				0
ENSG00000188234	AGAP4				0
ENSG00000176894	PXMP2				0
ENSG00000110047	EHD1				0
ENSG00000196361	ELAVL3				0
ENSG00000089199	CHGB				0
ENSG00000105063	PPP6R1				0
ENSG00000100665	SERPINA4	ENSGALG00000020388		Amniota	39.18
ENSG00000100665	SERPINA4	ENSGALG00000023070	SERPINA1	Amniota	40.2353
ENSG00000100665	SERPINA4	ENSGALG00000020390		Amniota	40.2685
ENSG00000204131	NHSL2	ENSGALG00000028506	NHSL2	Amniota	41.0194
ENSG00000136153	LMO7	ENSGALG00000016920	LMO7	Amniota	43.3354
ENSG00000100665	SERPINA4	ENSGALG00000010969	SERPINA4	Amniota	43.5294
ENSG00000198961	PJA2	ENSGALG00000000264		Tetrapoda	47.0333

ENSG00000113387	SUB1	ENSGALG00000027147		Amniota	48.8722
ENSG00000039523	FAM65A	ENSGALG00000030055		Amniota	51.9643
ENSG00000113555	PCDH12	ENSGALG00000002544	PCDH12	Amniota	52.9937
ENSG00000172375	C2CD2L	ENSGALG00000034014	C2CD2L	Amniota	53.2877
ENSG00000099290	FAM21A	ENSGALG00000005914	FAM21A	Amniota	54.566
ENSG00000130723	PRRC2B	ENSGALG00000038255		Chordata	54.721
ENSG00000144218	AFF3	ENSGALG00000031502		Amniota	57.4026
ENSG00000130881	LRP3	ENSGALG00000004853	LRP3	Sarcopterygii	57.6739
ENSG00000141699	FAM134C	ENSGALG00000003162	FAM134C	Amniota	61.2565
ENSG00000139631	CSAD	ENSGALG00000035850	CSAD	Amniota	61.3592
ENSG00000128159	TUBGCP6	ENSGALG00000032156	TUBGCP6	Amniota	61.6259
ENSG00000188643	S100A16	ENSGALG00000028924	S100A16	Amniota	61.9048
ENSG00000176871	WSB2	ENSGALG00000008151	WSB2	Amniota	63.7149
ENSG00000164985	PSIP1	ENSGALG00000015105	PSIP1	Amniota	64.7668
ENSG00000076864	RAP1GAP	ENSGALG00000014701	RAP1GAP	Euteleostomi	66.3669
ENSG00000126217	MCF2L	ENSGALG00000016834	MCF2L	Amniota	67.1962
ENSG00000143420	ENSA	ENSGALG00000039991		Amniota	68.8679
ENSG00000181163	NPM1	ENSGALG00000002197	NPM1	Amniota	70.068
ENSG00000188536	HBA2	ENSGALG00000043234	HBAA	Amniota	70.4225
ENSG00000078674	PCM1	ENSGALG00000013602	PCM1	Amniota	70.8508
ENSG00000187164	SHTN1	ENSGALG00000030276		Amniota	73.3227
ENSG00000122299	ZC3H7A	ENSGALG00000003143	ZC3H7A	Amniota	73.4295
ENSG00000005156	LIG3	ENSGALG00000002288	LIG3	Amniota	73.5119
ENSG00000078618	NRDC	ENSGALG00000010565	NRDC	Amniota	74.4607
ENSG00000057019	DCBLD2	ENSGALG00000015258	DCBLD2	Amniota	74.7041

ENSG00000113269	RNF130	ENSGALG00000005820	RNF130	Amniota	74.8815
ENSG00000115756	HPCAL1	ENSGALG00000016443	HPCAL1	Vertebrata	74.8899
ENSG00000018236	CNTN1	ENSGALG00000009523	CNTN1	Amniota	75.9531
ENSG00000198952	SMG5	ENSGALG00000029918	SMG5	Amniota	76.0724
ENSG00000153048	CARHSP1	ENSGALG00000042981	CARHSP1	Amniota	76.3158
ENSG00000108433	GOSR2	ENSGALG00000001107	GOSR2	Amniota	76.4151
ENSG00000138796	HADH	ENSGALG00000016124	HADH	Amniota	76.8254
ENSG00000144218	AFF3	ENSGALG00000016767		Amniota	77.0115
ENSG00000091136	LAMB1	ENSGALG00000007905	LAMB1	Amniota	78.0374
ENSG00000067248	DHX29	ENSGALG00000014700	DHX29	Amniota	78.1341
ENSG00000075240	GRAMD4	ENSGALG00000035169		Amniota	78.5047
ENSG00000141576	RNF157	ENSGALG00000002008	RNF157	Amniota	78.82
ENSG00000115241	PPM1G	ENSGALG00000016513	PPM1G	Amniota	79.0566
ENSG00000104969	SGTA	ENSGALG00000014067	SGTA	Amniota	79.2332
ENSG00000198522	GPN1	ENSGALG00000016505	GPN1	Amniota	79.4038
ENSG00000127022	CANX	ENSGALG00000032148	CANX	Amniota	80.3661
ENSG00000167770	OTUB1	ENSGALG00000038229	OTUB1	Amniota	80.3704
ENSG00000130723	PRRC2B	ENSGALG00000030256		Chordata	80.6748
ENSG00000075240	GRAMD4	ENSGALG00000034274		Amniota	80.7407
ENSG00000130723	PRRC2B	ENSGALG00000041505		Chordata	80.8824
ENSG00000172531	PPP1CA	ENSGALG00000040016	PPP1CA	Amniota	81.039
ENSG00000139289	PHLDA1	ENSGALG00000034788		Amniota	81.0458
ENSG00000163528	CHCHD4	ENSGALG00000006328	CHCHD4	Amniota	82.6087
ENSG00000126775	ATG14	ENSGALG00000012148	ATG14	Amniota	82.7935
ENSG00000143420	ENSA	ENSGALG00000037480	ENSA	Amniota	82.906

ENSG00000154118	JPH3	ENSGALG00000029424	JPH3	Amniota	83.1135
ENSG00000102893	PHKB	ENSGALG00000004004	PHKB	Amniota	84.1386
ENSG00000118971	CCND2	ENSGALG00000017283	CCND2	Sarcopterygii	84.5361
ENSG00000134982	APC	ENSGALG00000000220		Amniota	84.6127
ENSG00000119760	SUPT7L	ENSGALG00000010048	SUPT7L	Amniota	86.2319
ENSG00000113387	SUB1	ENSGALG00000034905		Amniota	86.5079
ENSG00000113387	SUB1	ENSGALG00000034827		Amniota	86.5079
ENSG00000114423	CBLB	ENSGALG00000015351	CBLB	Amniota	87.3922
ENSG00000008853	RHOBTB2	ENSGALG00000034020	RHOBTB2	Amniota	87.4339
ENSG00000071127	WDR1	ENSGALG00000014957	WDR1	Amniota	88.0131
ENSG00000161526	SAP30BP	ENSGALG00000002415	SAP30BP	Amniota	88.4244
ENSG00000075413	MARK3	ENSGALG00000011505	MARK3	Sarcopterygii	88.9029
ENSG00000113387	SUB1	ENSGALG00000003248	SUB1	Amniota	90.4762
ENSG00000101557	USP14	ENSGALG00000014920	USP14	Amniota	90.6504
ENSG00000166333	ILK	ENSGALG00000031642	ILK	Amniota	90.708
ENSG00000001631	KRIT1	ENSGALG00000042651	KRIT1	Amniota	92.1196
ENSG00000001631	KRIT1	ENSGALG00000033812		Amniota	92.1933
ENSG00000187109	NAP1L1	ENSGALG00000010238		Amniota	92.3664
ENSG00000140350	ANP32A	ENSGALG00000008054		Amniota	92.7419
ENSG00000126214	KLC1	ENSGALG00000011528		Tetrapoda	92.7786
ENSG00000087258	GNAO1	ENSGALG00000003163	GNAO1	Euteleostomi	92.9379
ENSG00000184432	COPB2	ENSGALG00000005357	COPB2	Amniota	92.9901
ENSG00000096384	HSP90AB1	ENSGALG00000010175	HSP90AB1	Amniota	94.3448
ENSG00000169727	GPS1	ENSGALG00000002827	GPS1	Amniota	94.4558
ENSG00000089157	RPLP0	ENSGALG00000023294	RPLP0	Amniota	94.6203

ENSG00000163466	ARPC2	ENSGALG00000011452	ARPC2	Amniota	94.6667
ENSG00000117360	PRPF3	ENSGALG00000028205	PRPF3	Amniota	94.883
ENSG00000011304	PTBP1	ENSGALG00000001962	PTBP1	Amniota	95.2471
ENSG00000119866	BCL11A	ENSGALG00000034048	BCL11A	Amniota	95.9799
ENSG00000116586	LAMTOR2	ENSGALG00000029881	LAMTOR2	Amniota	96
ENSG00000099956	SMARCB1	ENSGALG00000005983	SMARCB1	Amniota	96.3731
ENSG00000175792	RUVBL1	ENSGALG00000005931	RUVBL1	Amniota	96.4912
ENSG00000170759	KIF5B	ENSGALG00000007215	KIF5B	Amniota	96.5839
ENSG00000179456	ZBTB18	ENSGALG00000010705	ZBTB18	Amniota	97.3635
ENSG00000167658	EEF2	ENSGALG00000033884	EEF2	Amniota	97.4359
ENSG00000070087	PFN2	ENSGALG00000010410	PFN2	Amniota	98.5714
ENSG00000034713	GABARAPL2	ENSGALG00000003208	GABARAPL2	Amniota	99.1453
ENSG00000177889	UBE2N	ENSGALG00000030515	UBE2N	Amniota	100
ENSG00000113558	SKP1	ENSGALG00000006474	SKP1	Amniota	100
ENSG00000063245	EPN1				
ENSG00000100345	MYH9	ENSGALG00000043092		Amniota	92.8061
ENSG00000131504	DIAPH1				
ENSG00000089280	FUS	ENSGALG00000032219		Amniota	63.6538
ENSG00000119487	MAPKAP1	ENSGALG00000000977		Amniota	89.6552
ENSG00000160799	CCDC12	ENSGALG00000005487		Amniota	40.2174
ENSG00000163660	CCNL1	ENSGALG00000009699		Amniota	89.272
ENSG00000171206	TRIM8	ENSGALG00000036785		Amniota	85.2995
ENSG00000171206	TRIM8	ENSGALG00000036953		Amniota	77.7778
ENSG00000169567	HINT1	ENSGALG00000035863		Amniota	54.1667
ENSG00000169567	HINT1	ENSGALG00000041637		Amniota	30.7692

ENSG00000169567	HINT1	ENSGALG00000035489		Amniota	27.3333
ENSG00000169567	HINT1	ENSGALG00000035998		Amniota	26.6667
ENSG00000169567	HINT1	ENSGALG00000037690		Amniota	26.6667
ENSG00000169567	HINT1	ENSGALG00000039296		Amniota	35.7143
ENSG00000169567	HINT1	ENSGALG00000031712		Amniota	29.1971
ENSG00000169567	HINT1	ENSGALG00000029597		Amniota	41.7391
ENSG00000169567	HINT1	ENSGALG00000032641		Amniota	35.7143
ENSG00000169567	HINT1	ENSGALG00000000428		Amniota	82.5397
ENSG00000169567	HINT1	ENSGALG00000036558		Amniota	37.9845
ENSG00000100342	APOL1				
ENSG00000087365	SF3B2				
ENSG00000186918	ZNF395	ENSGALG00000034140		Amniota	63.7113

Table 8.10 - Orthology analysis of antigens identified from 384 scFv using Unipex protein array technology.

Human Ensembl Gene ID	Associated Gene Name	Chicken Ensembl Gene ID	Chicken Associated Gene Name	Last common ancestor with Chicken	Similarity Percentage
ENSG00000071655	MBD3	ENSGALG00000001101	MBD3	Amniota	75.3205
ENSG00000091640	SPAG7				
ENSG00000100380	ST13	ENSGALG00000012007	ST13	Amniota	81.9945
ENSG00000100417	PMM1	ENSGALG00000038599	PMM1	Amniota	65.4237
ENSG00000100442	FKBP3	ENSGALG00000012466	FKBP3	Amniota	65.9836
ENSG00000020256	ZFP64	ENSGALG00000038858	ZFP64	Vertebrata	73.5849
ENSG00000103266	STUB1	ENSGALG00000041432	STUB1	Amniota	83.4395
ENSG00000103495	MAZ				
ENSG00000104969	SGTA	ENSGALG00000014067	SGTA	Amniota	79.2332

ENSG00000086065	CHMP5	ENSGALG00000013160	CHMP5	Amniota	93.6073
ENSG00000109917	ZPR1	ENSGALG00000007101	ZPR1	Amniota	63.0901
ENSG00000089157	RPLP0	ENSGALG00000023294	RPLP0	Amniota	94.6203
ENSG00000117751	PPP1R8	ENSGALG00000000729	PPP1R8	Amniota	90.3955
ENSG00000118785	SPP1	ENSGALG00000010926	SPP1	Amniota	27.6515
ENSG00000119321	FKBP15	ENSGALG00000008838	FKBP15	Amniota	54.6392
ENSG00000124541	RRP36	ENSGALG00000028056	RRP36	Amniota	50
ENSG00000126705	AHDC1	ENSGALG00000042916	AHDC1	Amniota	62.5389
ENSG00000196453	ZNF777	ENSGALG00000038657	ZNF777	Amniota	55.3681
ENSG00000105429	MEGF8				
ENSG00000135482	ZC3H10	ENSGALG00000037941	ZC3H10	Amniota	83.2
ENSG00000138078	PREPL	ENSGALG00000009981	PREPL	Amniota	46.7213
ENSG00000065882	TBC1D1	ENSGALG00000039380	TBC1D1	Amniota	73.021
ENSG00000103043	VAC14	ENSGALG00000002458	VAC14	Amniota	85.2564
ENSG00000104131	EIF3J	ENSGALG00000008187	EIF3J	Amniota	80.2372
ENSG00000063046	EIF4B	ENSGALG00000043411		Amniota	79.6748
ENSG00000103126	AXIN1	ENSGALG00000029402	AXIN1	Amniota	74.1974
ENSG00000142002	DPP9	ENSGALG00000004200	DPP9	Amniota	87.2054
ENSG00000064666	CNN2	ENSGALG00000027777	CNN2	Amniota	84.918
ENSG00000115866	DARS	ENSGALG00000012355	DARS	Amniota	89.4632
ENSG00000118231	CRYGD				
ENSG00000091490	SEL1L3	ENSGALG00000039549	SEL1L3	Amniota	59.851
ENSG00000003147	ICA1	ENSGALG00000010708	ICA1	Amniota	51.3314
ENSG00000001629	ANKIB1	ENSGALG00000031917	ANKIB1	Amniota	87.66
ENSG00000140740	UQCRC2	ENSGALG00000002490	UQCRC2	Amniota	70.6783

ENSG00000156110	ADK	ENSGALG00000005057	ADK	Amniota	84.4011
ENSG00000157782	CABP1	ENSGALG00000007116	CABP1	Amniota	39.801
ENSG00000062598	ELMO2	ENSGALG00000007246	ELMO2	Amniota	95.5119
ENSG00000244045	TMEM199	ENSGALG00000023757	TMEM199	Amniota	67.0103
ENSG00000161010	C5orf45	ENSGALG00000035476		Amniota	23.9669
ENSG00000161526	SAP30BP	ENSGALG00000002415	SAP30BP	Amniota	88.4244
ENSG00000152795	HNRNPDL	ENSGALG00000023199	HNRNPDL	Amniota	93.6877
ENSG00000164603	BMT2	ENSGALG00000009440	C1H7ORF60	Amniota	87.0098
ENSG00000224016	RP11-728K20.1				
ENSG00000164929	BAALC				
ENSG00000165006	UBAP1	ENSGALG00000021378	UBAP1	Amniota	72.4409
ENSG00000165238	WNK2	ENSGALG00000038850	WNK2	Amniota	55.8849
ENSG00000011114	BTBD7	ENSGALG00000033360		Amniota	89.8795
ENSG00000011114	BTBD7	ENSGALG00000034311		Amniota	80.9683
ENSG00000166734	CASC4	ENSGALG00000008208	CASC4	Amniota	63.0979
ENSG00000166922	SCG5	ENSGALG00000009725	SCG5	Amniota	82.6923
ENSG00000167193	CRK	ENSGALG00000002656	CRK	Amniota	84.5902
ENSG00000167193	CRK	ENSGALG00000031595		Amniota	76.9517
ENSG00000164045	CDC25A	ENSGALG00000038469	CDC25A	Amniota	60.4563
ENSG00000171532	NEUROD2	ENSGALG00000039633	NEUROD2	Amniota	68.5851
ENSG00000171720	HDAC3	ENSGALG00000002634	HDAC3	Amniota	97.8972
ENSG00000170242	USP47	ENSGALG00000005569	USP47	Amniota	89.3169
ENSG00000169217	CD2BP2				
ENSG00000172239	PAIP1	ENSGALG00000014867	PAIP1	Euteleostomi	73.8397
ENSG00000106524	ANKMY2	ENSGALG00000010804	ANKMY2	Amniota	72.8261

ENSG00000174953	DHX36	ENSGALG00000043082	DHX36	Amniota	80.6484
ENSG00000138594	TMOD3	ENSGALG00000004710	TMOD3	Amniota	76.6304
ENSG00000167658	EEF2	ENSGALG00000033884	EEF2	Amniota	97.4359
ENSG00000141699	FAM134C	ENSGALG00000003162	FAM134C	Amniota	61.2565
ENSG00000155849	ELMO1	ENSGALG00000035492	ELMO1	Amniota	96.9739
ENSG00000196502	SULT1A1				
ENSG00000176973	FAM89B				
ENSG00000179218	CALR	ENSGALG00000040368	CALR	Amniota	75.0725
ENSG00000174996	KLC2				
ENSG00000116212	LRRC42	ENSGALG00000010743	LRRC42	Amniota	71.9715
ENSG00000114745	GORASP1	ENSGALG00000006033	GORASP1	Amniota	53.9906
ENSG00000138380	CARF	ENSGALG00000008474	CARF	Amniota	70.8108
ENSG00000111641	NOP2	ENSGALG00000037807		Amniota	68.8889
ENSG00000100644	HIF1A	ENSGALG00000011870	HIF1A	Amniota	79.0382
ENSG00000153207	AHCTF1	ENSGALG00000010630	AHCTF1	Amniota	57.231
ENSG00000171988	JMJD1C	ENSGALG00000002942	JMJD1C	Amniota	78.4895
ENSG00000181555	SETD2	ENSGALG00000042051	SETD2	Amniota	66.1508
ENSG00000184661	CDCA2				
ENSG00000064607	SUGP2	ENSGALG00000003047	SUGP2	Amniota	51.7958
ENSG00000151276	MAGI1	ENSGALG00000007431	MAGI1	Amniota	85.2524
ENSG00000182866	LCK	ENSGALG00000000466	LCK	Amniota	82.874
ENSG00000185049	NELFA	ENSGALG00000015699	NELFA	Amniota	84.0445
ENSG00000141391	PRELID3A	ENSGALG00000013830	SLMO1	Amniota	75.5208
ENSG00000106367	AP1S1	ENSGALG00000031292		Euteleostomi	40.099
ENSG00000122565	CBX3	ENSGALG00000011038	CBX3	Amniota	95.6522

ENSG00000119760	SUPT7L	ENSGALG00000010048	SUPT7L	Amniota	86.2319
ENSG00000096070	BRPF3	ENSGALG00000000776		Amniota	85.2217
ENSG00000111653	ING4	ENSGALG00000014457	ING4	Amniota	96.7871
ENSG00000148308	GTF3C5	ENSGALG00000003441	GTF3C5	Amniota	66.6667
ENSG00000105137	SYDE1	ENSGALG00000040312	SYDE1	Amniota	44.4609
ENSG00000140950	TLDC1	ENSGALG00000033137	TLDC1	Amniota	58.3333
ENSG00000067191	CACNB1	ENSGALG00000025788		Amniota	83.5616
ENSG00000101126	ADNP	ENSGALG00000007989	ADNP	Amniota	75.57
ENSG00000172273	HINFP	ENSGALG00000006806	HINFP	Amniota	65.5769
ENSG00000169062	UPF3A	ENSGALG00000016815	UPF3A	Amniota	59.3117
ENSG00000131473	ACLY	ENSGALG00000003475	ACLY	Amniota	93.3697
ENSG00000165283	STOML2	ENSGALG00000002064		Amniota	82.1782
ENSG00000165283	STOML2	ENSGALG00000029723		Amniota	82.1782
ENSG00000197694	SPTAN1	ENSGALG00000004719	SPTAN1	Amniota	95.6137
ENSG00000076924	XAB2				
ENSG00000161011	SQSTM1	ENSGALG00000035804	SQSTM1	Amniota	58.6288
ENSG00000168309	FAM107A	ENSGALG00000007139	FAM107A	Amniota	71.5278
ENSG00000142864	SERBP1	ENSGALG00000011229	SERBP1	Amniota	78.5377
ENSG00000171603	CLSTN1	ENSGALG00000002606	CLSTN1	Amniota	82.5953
ENSG00000181191	PJA1	ENSGALG00000000264		Tetrapoda	38.0608
ENSG00000198783	ZNF830	ENSGALG00000013156	ZNF830	Amniota	63.5616
ENSG00000173726	TOMM20	ENSGALG00000035299	TOMM20	Amniota	96.5753
ENSG00000196591	HDAC2	ENSGALG00000014991	HDAC2	Amniota	99.5902
ENSG00000130830	MPP1	ENSGALG00000005071	MPP1	Amniota	83.1197
ENSG00000134222	PSRC1				

ENSG00000101189	MRGBP	ENSGALG00000005531	MRGBP	Amniota	71.5084
ENSG00000101182	PSMA7	ENSGALG00000030411	PSMA7	Amniota	98.3871
ENSG00000136819	C9orf78	ENSGALG00000004152	C9orf78	Amniota	87.5433
ENSG00000116983	HPCAL4	ENSGALG00000003829	HPCAL4	Amniota	91.623
ENSG00000204256	BRD2	ENSGALG00000000156	BRD2	Amniota	82.1566
ENSG00000130956	HABP4	ENSGALG00000012628	HABP4	Amniota	55.1821
ENSG00000125817	CENPB				
ENSG00000126214	KLC1	ENSGALG00000011528		Tetrapoda	92.7786
ENSG00000084754	HADHA	ENSGALG00000016536	HADHA	Amniota	77.1429
ENSG00000211746	TRBV19	ENSGALG00000014750		Amniota	36.9128
ENSG00000211746	TRBV19	ENSGALG00000037804		Amniota	36.9128
ENSG00000211746	TRBV19	ENSGALG00000022907		Amniota	35.5705
ENSG00000211746	TRBV19	ENSGALG00000019249		Amniota	34.4371
ENSG00000104892	KLC3				
ENSG00000079246	XRCC5	ENSGALG00000011492	XRCC5	Amniota	62.8906
ENSG00000163938	GNL3	ENSGALG00000001613	GNL3	Amniota	50.7042
ENSG00000276219	ZNF676				
ENSG00000085644	ZNF213				
ENSG00000140320	BAHD1	ENSGALG00000033675	BAHD1	Amniota	52.2406
ENSG00000169714	CNBP	ENSGALG00000005001	CNBP	Amniota	100
ENSG00000230532	AC091133.1				
ENSG00000180764	PIPSL				
ENSG00000104969	SGTA	ENSGALG00000014067	SGTA	Amniota	79.2332
ENSG00000101384	JAG1	ENSGALG00000009020	JAG1	Amniota	88.3019
ENSG00000066777	ARFGEF1	ENSGALG00000034610	ARFGEF1	Amniota	94.7996

ENSG00000103126	AXIN1	ENSGALG00000029402	AXIN1	Amniota	74.1974
ENSG00000165006	UBAP1	ENSGALG00000021378	UBAP1	Amniota	72.4409
ENSG00000157514	TSC22D3	ENSGALG00000027069	TSC22D3	Euteleostomi	55.3957
ENSG00000179218	CALR	ENSGALG00000040368	CALR	Amniota	75.0725
ENSG00000068831	RASGRP2				
ENSG00000171345	KRT19	ENSGALG00000019719	KRT19	Amniota	72.104
ENSG00000158615	PPP1R15B	ENSGALG00000000611	PPP1R15B	Amniota	45.4874
ENSG00000204256	BRD2	ENSGALG00000000156	BRD2	Amniota	82.1566
ENSG00000079246	XRCC5	ENSGALG00000011492	XRCC5	Amniota	62.8906
ENSG00000180764	PIPSL				

Table 8.11 – DAVID analysis for enrichment of antigens profiles from chicken ABG055's serum-contained IgY after PCa cell line.

Annotation Cluster 1	**Enrichment Score: 1.0445965437307996**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
KEGG_PATHWAY	hsa04810:Regulation of actin cytoskeleton	6	5.405405405	0.025432213	1729, 5499, 55845, 324, 10109, 5217	3.447154472	0.917819446	0.917819446	24.7452881
KEGG_PATHWAY	hsa05131:Shigellosis	3	2.702702703	0.148483247	1729, 10109, 5217	4.282828283	0.999999831	0.925619913	83.03177593
UP_KEYWORDS	Actin-binding	5	4.504504505	0.194602654	9948, 1729, 4627, 10109, 5217	2.167826637	1	0.928661435	93.0962192
Annotation Cluster 2	**Enrichment Score: 0.8724685427782515**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR011990:Tetratricopeptide-like helical	6	5.405405405	0.029432266	2873, 29066, 4678, 6449, 3831, 23381	3.416883117	0.999855572	0.999855572	32.86004258
UP_SEQ_FEATURE	repeat:TPR 3	4	3.603603604	0.102910544	29066, 4678, 6449, 3831	3.507484307	1	0.99992286	78.42730548
UP_SEQ_FEATURE	repeat:TPR 1	4	3.603603604	0.124150271	29066, 4678, 6449, 3831	3.224622669	1	0.999934453	84.62059884
UP_SEQ_FEATURE	repeat:TPR 2	4	3.603603604	0.124150271	29066, 4678, 6449, 3831	3.224622669	1	0.999934453	84.62059884

UP_KEYWORDS	TPR repeat	4	3.603603604	0.144905924	29066, 4678, 6449, 3831	3.002601671	1	0.889601069	85.53657459
SMART	SM00028:SM00028	3	2.702702703	0.196096824	4678, 6449, 3831	3.617045455	0.999999938	0.999999938	89.94323216
INTERPRO	IPR019734:Tetratricopeptide repeat	3	2.702702703	0.259791986	4678, 6449, 3831	3.006857143	1	0.999985341	98.1897203
INTERPRO	IPR013026:Tetratricopeptide repeat-containing domain	3	2.702702703	0.304011693	4678, 6449, 3831	2.684693878	1	0.999978071	99.20385515
Annotation Cluster 3	**Enrichment Score: 0.6857768918067765**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0070911~global genome nucleotide-excision repair	3	2.702702703	0.154429091	2873, 7334, 3980	4.241142857	1	0.999963459	91.91238895
GOTERM_BP_DIRECT	GO:0006281~DNA repair	7	6.306306306	0.179283623	2873, 6598, 55611, 7334, 4869, 8607, 3980	1.822947368	1	0.999983815	94.82908273
GOTERM_BP_DIRECT	GO:0006289~nucleotide-excision repair	3	2.702702703	0.316518859	2873, 7334, 3980	2.604210526	1	0.999975671	99.66719439
Annotation Cluster 4	**Enrichment Score: 0.6790459420529562**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR018247:EF-Hand 1, calcium-binding site	4	3.603603604	0.098147118	140576, 3241, 10938, 4924	3.579591837	1	0.999999771	74.78234794
INTERPRO	IPR011992:EF-hand-like domain	5	4.504504505	0.10801354	140576, 3241, 10938, 868, 4924	2.723602484	1	0.999987352	78.22318489
INTERPRO	IPR002048:EF-hand domain	4	3.603603604	0.169658018	140576, 3241, 10938, 4924	2.784126984	1	0.999998941	91.62011086
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	7	6.306306306	0.274725151	140576, 3241, 10938, 868, 51294, 4924, 821	1.578923767	1	0.991715433	98.17917555
UP_SEQ_FEATURE	domain:EF-hand 1	3	2.702702703	0.297872234	140576, 3241, 4924	2.726271893	1	1	99.32243376
UP_SEQ_FEATURE	domain:EF-hand 2	3	2.702702703	0.297872234	140576, 3241, 4924	2.726271893	1	1	99.32243376
UP_KEYWORDS	Calcium	7	6.306306306	0.402522166	140576, 3241, 10938, 868, 51294, 4924, 821	1.359285891	1	0.973352923	99.82731937
Annotation Cluster 5	**Enrichment Score: 0.5746730748319002**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_MF_DIRECT	GO:0008017~microtubule binding	5	4.504504505	0.097375772	889, 11345, 324, 85378, 3799	2.825842697	0.999999998	0.98122106	72.13182873
UP_KEYWORDS	Microtubule	4	3.603603604	0.406906281	324, 85378, 3831, 3799	1.719438563	1	0.971007857	99.84233411

GOTERM_CC_DIRECT	GO:0005874~microtubule	4	3.603603604	0.476469897	889, 85378, 3831, 3799	1.543231962	1	0.996894763	99.97535469
Annotation Cluster 6	**Enrichment Score: 0.5595991624925843**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_MF_DIRECT	GO:0005525~GTP binding	6	5.405405405	0.1176349	23221, 10938, 1938, 11321, 3326, 2775	2.286363636	1	0.951919186	79.00302064
INTERPRO	IPR027417:P-loop containing nucleoside triphosphate hydrolase	10	9.009009009	0.215552396	23221, 10938, 4627, 1938, 1106, 8607, 11321, 54505, 2775, 3799	1.495948827	1	0.999965201	96.07409367
GOTERM_MF_DIRECT	GO:0003924~GTPase activity	4	3.603603604	0.25515727	5909, 1938, 11321, 2775	2.260674157	1	0.991455026	97.46217114
UP_KEYWORDS	GTP-binding	4	3.603603604	0.396213928	23221, 1938, 11321, 2775	1.749341843	1	0.975107746	99.80340053
UP_SEQ_FEATURE	nucleotide phosphate-binding region:GTP	3	2.702702703	0.621122784	23221, 1938, 2775	1.428047182	1	1	99.99988852
Annotation Cluster 7	**Enrichment Score: 0.531861615159944**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0007049~cell cycle	5	4.504504505	0.105063807	2873, 5499, 894, 4678, 3980	2.748888889	1	0.999975688	81.06531376
GOTERM_BP_DIRECT	GO:0051301~cell division	5	4.504504505	0.320827203	5499, 894, 8607, 2029, 3980	1.73006993	1	0.99996859	99.69729589
UP_KEYWORDS	Cell division	5	4.504504505	0.432398962	5499, 894, 8607, 2029, 3980	1.479222882	1	0.975311398	99.90836323
UP_KEYWORDS	Cell cycle	7	6.306306306	0.51161654	6598, 5499, 894, 4678, 8607, 2029, 3980	1.213982917	1	0.978874764	99.98568677
Annotation Cluster 8	**Enrichment Score: 0.46151015291701464**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_CC_DIRECT	GO:0000139~Golgi membrane	7	6.306306306	0.182312529	9867, 9570, 5909, 9276, 22870, 11345, 8677	1.814503205	1	0.964400213	92.45282783
UP_KEYWORDS	Protein transport	7	6.306306306	0.445275688	9570, 9276, 11345, 4678, 131474, 8677, 387680	1.29909611	1	0.972529224	99.93097996
UP_KEYWORDS	Golgi apparatus	7	6.306306306	0.508187576	9867, 9570, 5909, 9276, 11345, 8677, 4924	1.21818355	1	0.98046164	99.98439484
Annotation Cluster 9	**Enrichment Score: 0.43626215503319826**								

Category	Term	Count	%	P-Value	Genes	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_DIRECT	GO:0016567~protein ubiquitination	8	7.207207207	0.018518827	9867, 7334, 4008, 868, 55819, 55884, 6500, 81603	2.910588235	0.999996362	0.999996362	24.4398014
GOTERM_MF_DIRECT	GO:0004842~ubiquitin-protein transferase activity	6	5.405405405	0.077490849	9867, 7334, 4008, 868, 55819, 6500	2.601724138	0.99999984	0.979997784	63.42921348
UP_KEYWORDS	Ubl conjugation pathway	9	8.108108108	0.161699688	9867, 55611, 7334, 9097, 868, 55819, 55884, 6500, 81603	1.676452599	1	0.900293447	88.67932837
GOTERM_MF_DIRECT	GO:0016874~ligase activity	4	3.603603604	0.395684031	9867, 868, 55819, 81603	1.749565217	1	0.996809659	99.81291023
INTERPRO	IPR001841:Zinc finger, RING-type	4	3.603603604	0.435502975	9867, 868, 55819, 81603	1.643091335	1	0.999999792	99.95121989
SMART	SM00184:SM00184	4	3.603603604	0.45223374	9867, 868, 55819, 81603	1.594290008	1	1	99.8224844
GOTERM_MF_DIRECT	GO:0061630~ubiquitin protein ligase activity	3	2.702702703	0.456987457	9867, 7334, 6500	1.934615385	1	0.998039941	99.95071672
UP_SEQ_FEATURE	zinc finger region:RING-type	3	2.702702703	0.467215199	868, 55819, 81603	1.898037394	1	1	99.98625576
UP_KEYWORDS	Ligase	5	4.504504505	0.467656524	9867, 868, 55819, 3980, 81603	1.412740955	1	0.975809259	99.95849831
INTERPRO	IPR013083:Zinc finger, RING/FYVE/PHD-type	4	3.603603604	0.738517131	9867, 868, 55819, 81603	1.044047619	1	1	99.9999983
GOTERM_MF_DIRECT	GO:0008270~zinc ion binding	8	7.207207207	0.754343324	9867, 4008, 868, 55819, 4898, 2521, 3980, 81603	0.935813953	1	0.999981424	99.99999751
UP_KEYWORDS	Metal-binding	21	18.91891892	0.86215107	140576, 3241, 53335, 55893, 868, 55819, 4898, 821, 3980, 2775, 10472, 9867, 5496, 10938, 29066, 5499, 4008, 3040, 2521, 4924, 81603	0.86147238	1	0.998459388	100
UP_KEYWORDS	Zinc-finger	10	9.009009009	0.942355954	10472, 9867, 29066, 53335, 55893, 868, 55819, 2521, 3980, 81603	0.723648604	1	0.999633638	100
UP_KEYWORDS	Zinc	12	10.81081081	0.954944051	10472, 9867, 29066, 53335, 4008, 55893, 868, 55819, 4898, 2521, 3980, 81603	0.718479685	1	0.999789654	100
Annotation Cluster 10	**Enrichment Score: 0.40615285408166446**								

Category	Term	Count	%	P-Value	Genes	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_DIRECT	GO:0006281~DNA repair	7	6.306306306	0.179283623	2873, 6598, 55611, 7334, 4869, 8607, 3980	1.822947368	1	0.999983815	94.82908273
UP_KEYWORDS	DNA repair	4	3.603603604	0.529494735	55611, 7334, 8607, 3980	1.42676817	1	0.978346099	99.99097001
UP_KEYWORDS	DNA damage	4	3.603603604	0.636999619	55611, 7334, 8607, 3980	1.219238254	1	0.990304395	99.99963339
Annotation Cluster 11	**Enrichment Score: 0.3853378594096214**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_CC_DIRECT	GO:0000139~Golgi membrane	7	6.306306306	0.182312529	9867, 9570, 5909, 9276, 22870, 11345, 8677	1.814503205	1	0.964400213	92.45282783
GOTERM_BP_DIRECT	GO:0018279~protein N-linked glycosylation via asparagine	4	3.603603604	0.28488445	9570, 9276, 22870, 821	2.128172043	1	0.99997741	99.34418365
GOTERM_BP_DIRECT	GO:0006888~ER to Golgi vesicle-mediated transport	3	2.702702703	0.403656252	9570, 9276, 22870	2.151304348	1	0.999995754	99.95692692
GOTERM_BP_DIRECT	GO:0043687~post-translational protein modification	5	4.504504505	0.501751423	9570, 9276, 1938, 22870, 821	1.351912568	1	0.99999928	99.9970889
GOTERM_BP_DIRECT	GO:0044267~cellular protein metabolic process	8	7.207207207	0.629584017	9570, 9276, 1938, 22870, 131474, 821, 6175, 3799	1.058395722	1	0.99999981	99.99996582
GOTERM_BP_DIRECT	GO:0061024~membrane organization	3	2.702702703	0.736094535	9570, 9276, 22870	1.150697674	1	0.999999982	99.99999979
Annotation Cluster 12	**Enrichment Score: 0.35320745625047356**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0045087~innate immune response	10	9.009009009	0.188681675	8542, 7334, 55845, 3326, 28956, 10109, 4140, 79109, 6500, 81603	1.54625	1	0.999954902	95.64909609
GOTERM_BP_DIRECT	GO:0038095~Fc-epsilon receptor signaling pathway	5	4.504504505	0.293503748	7334, 28956, 4140, 79109, 6500	1.805839416	1	0.999974607	99.45320058
GOTERM_BP_DIRECT	GO:0048010~vascular endothelial growth factor receptor signaling pathway	4	3.603603604	0.459336562	55845, 28956, 4140, 79109	1.58336	1	0.99999831	99.99009182
GOTERM_BP_DIRECT	GO:0007173~epidermal growth factor receptor signaling pathway	4	3.603603604	0.559480636	28956, 4140, 79109, 29924	1.364965517	1	0.999999643	99.99954053

GOTERM_BP_DIRECT	GO:0008543~fibroblast growth factor receptor signaling pathway	3	2.702702703	0.728108999	28956, 4140, 79109	1.168818898	1	0.999999987	99.99999967
GOTERM_BP_DIRECT	GO:0048011~neurotrophin TRK receptor signaling pathway	4	3.603603604	0.733329935	23263, 28956, 4140, 79109	1.052765957	1	0.999999986	99.99999975
Annotation Cluster 13	**Enrichment Score: 0.28849294798980385**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0015031~protein transport	5	4.504504505	0.36674681	9570, 4627, 11345, 4678, 387680	1.616993464	1	0.999992294	99.89402275
UP_KEYWORDS	Protein transport	7	6.306306306	0.445275688	9570, 9276, 11345, 4678, 131474, 8677, 387680	1.29909611	1	0.972529224	99.93097996
UP_KEYWORDS	Transport	9	8.108108108	0.834691187	8542, 9570, 9276, 11345, 3040, 4678, 131474, 8677, 387680	0.852433525	1	0.998433025	99.99999998
Annotation Cluster 14	**Enrichment Score: 0.27385153958746156**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR027417:P-loop containing nucleoside triphosphate hydrolase	10	9.009009009	0.215552396	23221, 10938, 4627, 1938, 1106, 8607, 11321, 54505, 2775, 3799	1.495948827	1	0.999965201	96.07409367
UP_KEYWORDS	Nucleotide-binding	16	14.41441441	0.382829533	23221, 3094, 1938, 1106, 8607, 3611, 3326, 2775, 3980, 7334, 10938, 4627, 11321, 54505, 4140, 3799	1.166227895	1	0.978504185	99.74225297
UP_KEYWORDS	ATP-binding	11	9.90990991	0.677076442	10938, 7334, 4627, 1106, 8607, 3611, 54505, 3326, 4140, 3980, 3799	0.987909567	1	0.991856337	99.99991357
UP_SEQ_FEATURE	nucleotide phosphate-binding region:ATP	8	7.207207207	0.689496705	10938, 4627, 1106, 8607, 3611, 54505, 4140, 3799	0.999633028	1	1	99.99999329
GOTERM_MF_DIRECT	GO:0005524~ATP binding	11	9.90990991	0.744157389	10938, 7334, 4627, 1106, 8607, 3611, 54505, 3326, 4140, 3980, 3799	0.934628378	1	0.999989852	99.99999586

UP_SEQ_FEATURE	binding site:ATP	4	3.603603604	0.793410196	10938, 3611, 4140, 3980	0.952031455	1	1	99.99999998
Annotation Cluster 15	**Enrichment Score: 0.17417412708837043**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR020472:G-protein beta WD-40 repeat	3	2.702702703	0.208414542	9948, 9276, 55884	3.496345515	1	0.99999017	95.57003126
UP_SEQ_FEATURE	repeat:WD 5	3	2.702702703	0.675243296	9948, 9276, 55884	1.292628915	1	1	99.99998736
UP_SEQ_FEATURE	repeat:WD 4	3	2.702702703	0.714763773	9948, 9276, 55884	1.199559633	1	1	99.99999798
UP_SEQ_FEATURE	repeat:WD 3	3	2.702702703	0.742676948	9948, 9276, 55884	1.135946622	1	1	99.99999953
INTERPRO	IPR001680:WD40 repeat	3	2.702702703	0.745100685	9948, 9276, 55884	1.130397422	1	1	99.99999879
UP_SEQ_FEATURE	repeat:WD 2	3	2.702702703	0.753914848	9948, 9276, 55884	1.110703364	1	1	99.99999975
UP_SEQ_FEATURE	repeat:WD 1	3	2.702702703	0.753914848	9948, 9276, 55884	1.110703364	1	1	99.99999975
UP_KEYWORDS	WD repeat	3	2.702702703	0.754470372	9948, 9276, 55884	1.109417161	1	0.99669055	99.99999707
SMART	SM00320:SM00320	3	2.702702703	0.760173629	9948, 9276, 55884	1.09607438	1	1	99.99997017
INTERPRO	IPR017986:WD40-repeat-containing domain	3	2.702702703	0.799745679	9948, 9276, 55884	1.009012464	1	1	99.99999995
INTERPRO	IPR015943:WD40/YVTN repeat-like-containing domain	3	2.702702703	0.836310173	9948, 9276, 55884	0.928042328	1	1	100
Annotation Cluster 16	**Enrichment Score: 0.12779060569133233**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_SEQ_FEATURE	topological domain:Cytoplasmic	12	10.81081081	0.616415718	641700, 131566, 57338, 9570, 93109, 4037, 8677, 51294, 55819, 821, 112476, 5827	1.028781847	1	1	99.99986729
GOTERM_CC_DIRECT	GO:0016021~integral component of membrane	12	10.81081081	0.652257069	641700, 57338, 9570, 93109, 9854, 162427, 4037, 8677, 23151, 55819, 112476, 5827	1.002065049	1	0.999404466	99.99987099
UP_SEQ_FEATURE	transmembrane region	15	13.51351351	0.786158695	641700, 9570, 93109, 55819, 23151, 8677, 821, 131566, 57338,	0.902195873	1	1	99.99999997

					9854, 162427, 4037, 51294, 112476, 5827				
UP_KEYWORDS	Transmembrane helix	15	13.51351351	0.804971037	641700, 9570, 93109, 55819, 23151, 8677, 821, 131566, 57338, 9854, 162427, 4037, 51294, 112476, 5827	0.888579116	1	0.997767777	99.99999983
UP_KEYWORDS	Transmembrane	15	13.51351351	0.811996437	641700, 9570, 93109, 55819, 23151, 8677, 821, 131566, 57338, 9854, 162427, 4037, 51294, 112476, 5827	0.883376662	1	0.997794801	99.99999989
UP_SEQ_FEATURE	topological domain:Extracellular	7	6.306306306	0.828169566	641700, 131566, 93109, 4037, 51294, 55819, 112476	0.859635282	1	1	100
Annotation Cluster 17	**Enrichment Score: 0.09220016493584501**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_SEQ_FEATURE	signal peptide	13	11.71171171	0.638585379	131566, 641700, 8542, 1272, 4037, 1114, 55819, 51294, 3912, 5267, 4898, 4924, 821	1.007382121	1	1	99.99994275
UP_KEYWORDS	Signal	13	11.71171171	0.668618506	131566, 641700, 8542, 1272, 4037, 1114, 55819, 51294, 3912, 5267, 4898, 4924, 821	0.986148588	1	0.991871664	99.99988106
UP_KEYWORDS	Glycoprotein	14	12.61261261	0.803746522	131566, 641700, 8542, 1272, 3040, 4037, 1114, 55819, 51294, 3912, 5267, 3326, 4924, 112476	0.887906799	1	0.997986896	99.99999982
UP_KEYWORDS	Disulfide bond	9	8.108108108	0.869912871	131566, 1272, 7334, 4869, 131474, 4037, 1114, 3912, 821	0.81410468	1	0.998210939	100
UP_SEQ_FEATURE	disulfide bond	6	5.405405405	0.956028316	131566, 1272, 4037, 1114, 3912, 821	0.654781459	1	1	100
UP_SEQ_FEATURE	glycosylation site:N-linked (GlcNAc...)	9	8.108108108	0.980248144	8542, 131566, 1272, 4037, 51294, 55819, 3912, 5267, 112476	0.62651095	1	1	100
Annotation Cluster 18	**Enrichment Score: 0.07538879733024337**								

Category	Term	Count	%	P-Value	Genes	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_DIRECT	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	7	6.306306306	0.770133945	7392, 23263, 11168, 6598, 10923, 53335, 57018	0.926096257	1	0.999999994	99.99999997
GOTERM_BP_DIRECT	GO:0006355~regulation of transcription, DNA-templated	8	7.207207207	0.807351566	11168, 3094, 9913, 55893, 29115, 23589, 3899, 8125	0.881603563	1	0.999999997	100
GOTERM_BP_DIRECT	GO:0006351~transcription, DNA-templated	12	10.81081081	0.830617589	10472, 11168, 3094, 6598, 1106, 9913, 55893, 29115, 8607, 3899, 57018, 8125	0.863023256	1	0.999999998	100
UP_KEYWORDS	Transcription regulation	15	13.51351351	0.842014703	6598, 3094, 1106, 53335, 9913, 29115, 55893, 8607, 3899, 8125, 7392, 10472, 11168, 10923, 57018	0.86020943	1	0.99848732	99.99999999
UP_KEYWORDS	Transcription	15	13.51351351	0.874644182	6598, 3094, 1106, 53335, 9913, 29115, 55893, 8607, 3899, 8125, 7392, 10472, 11168, 10923, 57018	0.832675132	1	0.998224418	100
UP_KEYWORDS	DNA-binding	10	9.009009009	0.92786768	7392, 10472, 11168, 6598, 10923, 1106, 55893, 3899, 4924, 2521	0.746195519	1	0.999517898	100
Annotation Cluster 19	**Enrichment Score: 0.018697038281415633**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR015880:Zinc finger, C2H2-like	4	3.603603604	0.929217013	10472, 29066, 53335, 55893	0.698456944	1	1	100
SMART	SM00355:SM00355	4	3.603603604	0.941870157	10472, 29066, 53335, 55893	0.672157111	1	1	100
INTERPRO	IPR007087:Zinc finger, C2H2	3	2.702702703	0.971908111	10472, 53335, 55893	0.563081862	1	1	100
GOTERM_MF_DIRECT	GO:0046872~metal ion binding	7	6.306306306	0.989642758	10472, 5496, 5499, 29066, 53335, 55893, 2775	0.556240126	1	1	100

Table 8.12 – DAVID analysis for enrichment in the antigen profile of 384 scFv.

Annotation Cluster 1	**Enrichment Score: 1.7316316983152071**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	Myristate	8	5.633802817	0.002827444	79870, 55727, 54467, 64689, 9478, 51440, 57707, 3932	4.151457781	0.441934245	0.252962012	3.503264282
UP_SEQ_FEATURE	lipid moiety-binding region:N-myristoyl glycine	5	3.521126761	0.028281011	79870, 64689, 9478, 51440, 3932	4.260243979	0.999999985	0.972921864	34.7174728
UP_KEYWORDS	Lipoprotein	10	7.042253521	0.079839156	79870, 55727, 4354, 30968, 54467, 64689, 9478, 51440, 57707, 3932	1.887026264	0.999999964	0.732465743	64.9346739
Annotation Cluster 2	**Enrichment Score: 1.540624342482631**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0006355~regulation of transcription, DNA-templated	22	15.49295775	0.003832371	811, 57473, 55734, 140467, 9987, 9913, 29115, 3091, 25988, 7760, 27153, 1831, 5511, 25909, 22927, 1059, 29072, 23394, 2969, 7555, 148213, 221037	1.924134762	0.969036877	0.969036877	5.813618897
UP_KEYWORDS	Transcription	34	23.94366197	0.015092827	56949, 79800, 170506, 9987, 5916, 25988, 55257, 8841, 5511, 4761, 22927, 6046, 7520, 23394, 4150, 22893, 148213, 221037, 9328, 57473, 55734, 140467, 3091, 29115, 9913, 53615, 7760, 27153, 23741, 29072, 3066, 2969, 7555, 11335	1.480045103	0.956405482	0.64805298	17.43112616
GOTERM_BP_DIRECT	GO:0006351~transcription, DNA-templated	27	19.01408451	0.020649006	56949, 170506, 9987, 25988, 55257, 8841, 5511, 22927, 6046, 7520, 23394, 22893, 148213, 221037, 9328, 57473, 55734, 140467, 29115, 9913, 53615, 7760, 27153, 23741, 3066, 7555, 11335	1.541112957	0.999999994	0.998153251	27.78108989
UP_KEYWORDS	Transcription regulation	32	22.53521127	0.027474166	79800, 170506, 9987, 5916, 25988, 55257, 8841, 5511, 4761, 22927, 6046, 7520, 23394, 4150,	1.439045799	0.996781628	0.559497852	29.59238723

					22893, 148213, 221037, 57473, 55734, 140467, 3091, 29115, 9913, 53615, 7760, 27153, 23741, 29072, 3066, 2969, 7555, 11335				
UP_KEYWORDS	DNA-binding	24	16.90140845	0.077202979	9328, 57473, 79800, 55734, 170506, 140467, 9987, 5916, 3091, 25988, 27245, 53615, 7760, 27153, 5511, 25909, 4761, 1059, 7520, 4150, 23394, 2969, 7555, 148213	1.404350703	0.999999935	0.777907449	63.64820363
GOTERM_MF_DIRECT	GO:0003677~DNA binding	18	12.67605634	0.22519468	811, 9328, 57473, 79800, 55734, 140467, 9987, 5916, 25988, 27245, 53615, 27153, 5511, 25909, 7520, 23394, 2969, 148213	1.282272798	1	0.998467177	96.39494955
Annotation Cluster 3	**Enrichment Score: 1.3977097935341742**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR019734:Tetratricopeptide repeat	7	4.929577465	0.001372917	56949, 10273, 6767, 147700, 6449, 64837, 3831	5.580909091	0.364519498	0.364519498	1.845318076
INTERPRO	IPR013026:Tetratricopeptide repeat-containing domain	7	4.929577465	0.002483664	56949, 10273, 6767, 147700, 6449, 64837, 3831	4.982954545	0.559845069	0.336558269	3.315121135
SMART	SM00028:SM00028	6	4.225352113	0.002919899	10273, 6767, 147700, 6449, 64837, 3831	5.894444444	0.231392838	0.231392838	3.131063957
INTERPRO	IPR002151:Kinesin light chain	3	2.112676056	0.00357143	147700, 64837, 3831	29.89772727	0.692930562	0.325349472	4.734733516
INTERPRO	IPR011990:Tetratricopeptide-like helical	8	5.633802817	0.0060074	56949, 10273, 6767, 23231, 147700, 6449, 64837, 3831	3.623966942	0.863088529	0.391711006	7.844107742
GOTERM_CC_DIRECT	GO:0035253~ciliary rootlet	3	2.112676056	0.009186064	147700, 64837, 3831	19.31343284	0.858639993	0.858639993	11.02127604
GOTERM_BP_DIRECT	GO:0008088~axon cargo transport	3	2.112676056	0.012235467	147700, 64837, 3831	16.82993197	0.999985501	0.996192181	17.47210197
UP_SEQ_FEATURE	repeat:TPR 3	6	4.225352113	0.014136501	10273, 6767, 147700, 6449, 64837, 3831	4.125709958	0.99987095	0.988639965	19.07385883
UP_SEQ_FEATURE	repeat:TPR 2	6	4.225352113	0.019769355	10273, 6767, 147700, 6449, 64837, 3831	3.792991413	0.999996489	0.98480053	25.68131403
UP_SEQ_FEATURE	repeat:TPR 1	6	4.225352113	0.019769355	10273, 6767, 147700, 6449, 64837, 3831	3.792991413	0.999996489	0.98480053	25.68131403
UP_KEYWORDS	TPR repeat	6	4.225352113	0.026089728	10273, 6767, 147700, 6449, 64837, 3831	3.531837217	0.995685773	0.596524503	28.31957527

GOTERM_CC_DIRECT	GO:0005871~kinesin complex	3	2.112676056	0.134118171	147700, 64837, 3831	4.635223881	1	0.993831142	83.83284581
UP_SEQ_FEATURE	repeat:TPR 5	3	2.112676056	0.148503615	147700, 64837, 3831	4.354916067	1	0.999959403	90.83371717
GOTERM_MF_DIRECT	GO:0003777~microtubule motor activity	3	2.112676056	0.1624634	147700, 64837, 3831	4.113958561	1	0.998392549	90.06316725
KEGG_PATHWAY	hsa05132:Salmonella infection	3	2.112676056	0.197105279	147700, 64837, 3831	3.581081081	1	0.999988978	91.41061373
UP_SEQ_FEATURE	repeat:TPR 4	3	2.112676056	0.259654838	147700, 64837, 3831	3.014941893	1	0.999999518	98.85399033
UP_KEYWORDS	Motor protein	3	2.112676056	0.350841877	147700, 64837, 3831	2.414623403	1	0.982504825	99.56686125
UP_KEYWORDS	Microtubule	3	2.112676056	0.799565242	147700, 64837, 3831	1.011252536	1	0.998063961	99.99999984
GOTERM_BP_DIRECT	GO:0008152~metabolic process	3	2.112676056	0.812895809	147700, 64837, 3831	0.981746032	1	1	100
GOTERM_CC_DIRECT	GO:0005874~microtubule	3	2.112676056	0.850254786	147700, 64837, 3831	0.898299202	1	0.999997707	100
Annotation Cluster 4	**Enrichment Score: 0.9731384230997453**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR008145:Guanylate kinase/L-type calcium channel	3	2.112676056	0.040097367	782, 4354, 9223	9.199300699	0.999998636	0.932858033	42.58146829
SMART	SM00072:SM00072	3	2.112676056	0.040670362	782, 4354, 9223	9.068376068	0.976170904	0.845633241	36.34493315
INTERPRO	IPR027417:P-loop containing nucleoside triphosphate hydrolase	8	5.633802817	0.738228539	782, 170506, 4354, 1938, 6817, 9557, 26354, 9223	0.951967436	1	1	99.99999872
Annotation Cluster 5	**Enrichment Score: 0.7534744365915169**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_SEQ_FEATURE	calcium-binding region:2	4	2.816901408	0.066128495	10235, 9478, 51440, 6709	4.237215633	1	0.995388682	63.83161803
UP_SEQ_FEATURE	calcium-binding region:1	4	2.816901408	0.079675591	10235, 9478, 51440, 6709	3.91942446	1	0.996980935	70.89313633
SMART	SM00054:SM00054	4	2.816901408	0.123859564	10235, 9478, 51440, 6709	3.207860922	0.999993214	0.981067336	76.27128472
UP_SEQ_FEATURE	domain:EF-hand 3	3	2.112676056	0.148503615	9478, 51440, 6709	4.354916067	1	0.999959403	90.83371717

UP_SEQ_FEATURE	domain:EF-hand 1	4	2.816901408	0.162169124	10235, 9478, 51440, 6709	2.850490517	1	0.999959641	92.79311298
UP_SEQ_FEATURE	domain:EF-hand 2	4	2.816901408	0.162169124	10235, 9478, 51440, 6709	2.850490517	1	0.999959641	92.79311298
INTERPRO	IPR018247:EF-Hand 1, calcium-binding site	4	2.816901408	0.162398672	10235, 9478, 51440, 6709	2.847402597	1	0.999941514	90.95081404
INTERPRO	IPR002048:EF-hand domain	4	2.816901408	0.266205243	10235, 9478, 51440, 6709	2.214646465	1	0.999988222	98.49484563
UP_KEYWORDS	Calcium	9	6.338028169	0.327379183	811, 782, 182, 10235, 22883, 1954, 9478, 51440, 6709	1.370461932	1	0.983171937	99.32261814
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	8	5.633802817	0.354247805	811, 182, 10235, 22883, 1954, 9478, 51440, 6709	1.377468935	1	0.997911576	99.66394295
INTERPRO	IPR011992:EF-hand-like domain	4	2.816901408	0.403001852	10235, 9478, 51440, 6709	1.733201581	1	0.99999981	99.90820001
Annotation Cluster 6	**Enrichment Score: 0.5436218912736942**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
INTERPRO	IPR001452:Src homology-3 domain	5	3.521126761	0.265047866	1398, 782, 4354, 3932, 6709	1.898268398	1	0.999996958	98.46234037
UP_KEYWORDS	SH3 domain	5	3.521126761	0.265971674	1398, 782, 4354, 3932, 6709	1.896098506	1	0.976408511	97.96434172
SMART	SM00326:SM00326	5	3.521126761	0.270021594	1398, 782, 4354, 3932, 6709	1.871252205	1	0.999159621	96.74174876
UP_SEQ_FEATURE	domain:SH3	4	2.816901408	0.3515234	782, 4354, 3932, 6709	1.888879258	1	0.999998786	99.84008896
Annotation Cluster 7	**Enrichment Score: 0.4065845884431345**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_MF_DIRECT	GO:0019899~enzyme binding	7	4.929577465	0.198984547	8312, 3091, 10273, 3066, 25988, 8841, 11335	1.768280434	1	0.999127836	94.43996822
KEGG_PATHWAY	hsa04919:Thyroid hormone signaling pathway	3	2.112676056	0.346163558	3091, 3066, 8841	2.409090909	1	0.999366911	99.13560845
GOTERM_MF_DIRECT	GO:0008134~transcription factor binding	3	2.112676056	0.875277042	3091, 3066, 8841	0.840809049	1	0.999999824	100
Annotation Cluster 8	**Enrichment Score: 0.3414266921752766**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**

GOTERM_BP_DIRECT	GO:0044255~cellular lipid metabolic process	4	2.816901408	0.232397259	3030, 65985, 47, 8841	2.37999038	1	0.999988782	98.38457579
UP_KEYWORDS	Lipid metabolism	4	2.816901408	0.436923215	3030, 65985, 47, 6817	1.643285372	1	0.98943844	99.92780691
GOTERM_BP_DIRECT	GO:0044281~small molecule metabolic process	11	7.746478873	0.931287558	5688, 132, 1615, 3030, 55697, 65985, 47, 6817, 8841, 6175, 7385	0.74734992	1	1	100
Annotation Cluster 9	**Enrichment Score: 0.32345774339130606**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
KEGG_PATHWAY	hsa05131:Shigellosis	3	2.112676056	0.164956895	1398, 63916, 9844	4.015151515	0.999999993	0.999999993	86.67687683
KEGG_PATHWAY	hsa05100:Bacterial invasion of epithelial cells	3	2.112676056	0.238273252	1398, 63916, 9844	3.154761905	1	0.99915515	95.23168371
GOTERM_BP_DIRECT	GO:0038096~Fc-gamma receptor signaling pathway involved in phagocytosis	3	2.112676056	0.277562473	1398, 63916, 9844	2.87340302	1	0.99999722	99.37269674
GOTERM_BP_DIRECT	GO:0048011~neurotrophin TRK receptor signaling pathway	7	4.929577465	0.338613987	1398, 5688, 8878, 3066, 8841, 3932, 6709	1.462174941	1	0.999999437	99.84175012
GOTERM_BP_DIRECT	GO:0048010~vascular endothelial growth factor receptor signaling pathway	5	3.521126761	0.388900284	1398, 5688, 63916, 9844, 6709	1.570793651	1	0.999999647	99.95390639
GOTERM_BP_DIRECT	GO:0006915~apoptotic process	7	4.929577465	0.642224572	5688, 63916, 9844, 51147, 8312, 29115, 6709	1.06546081	1	0.999999967	99.99989111
GOTERM_BP_DIRECT	GO:0000186~activation of MAPKK activity	3	2.112676056	0.657409377	1398, 5688, 6709	1.338744589	1	0.999999978	99.99999446
GOTERM_BP_DIRECT	GO:0045087~innate immune response	8	5.633802817	0.708300607	1398, 5688, 63916, 170506, 9844, 7520, 3932, 6709	0.981746032	1	0.999999991	99.99999955
GOTERM_BP_DIRECT	GO:0008543~fibroblast growth factor receptor signaling pathway	3	2.112676056	0.837199343	5688, 3932, 6709	0.927634046	1	1	100
GOTERM_BP_DIRECT	GO:0038095~Fc-epsilon receptor signaling pathway	3	2.112676056	0.867015124	5688, 3932, 6709	0.859923531	1	1	100
GOTERM_BP_DIRECT	GO:0007173~epidermal growth factor receptor signaling pathway	3	2.112676056	0.887195188	5688, 3932, 6709	0.812479475	1	1	100
Annotation Cluster 10	**Enrichment Score: 0.3130858666840036**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**

UP_KEYWORDS	Zinc-finger	22	15.49295775	0.21448952	57473, 8882, 55734, 10235, 140467, 91603, 5916, 54467, 25988, 27154, 7760, 27153, 51147, 57037, 8878, 64219, 4150, 23394, 84872, 7555, 148213, 221037	1.248423995	1	0.9636849	95.21942909
UP_SEQ_FEATURE	zinc finger region:C2H2-type 2	8	5.633802817	0.282285135	57473, 55734, 140467, 25988, 4150, 148213, 7760, 27153	1.49311408	1	0.999999089	99.27759814
UP_SEQ_FEATURE	zinc finger region:C2H2-type 3	8	5.633802817	0.306577193	57473, 55734, 140467, 25988, 23394, 4150, 7760, 27153	1.451638689	1	0.999998691	99.56699207
INTERPRO	IPR007087:Zinc finger, C2H2	9	6.338028169	0.346670076	57473, 55734, 140467, 25988, 23394, 4150, 148213, 7760, 27153	1.34371808	1	0.999999207	99.6883028
UP_SEQ_FEATURE	zinc finger region:C2H2-type 1	7	4.929577465	0.359235068	55734, 140467, 25988, 23394, 4150, 7760, 27153	1.428956835	1	0.999998378	99.86614082
UP_SEQ_FEATURE	zinc finger region:C2H2-type 5	7	4.929577465	0.363808451	57473, 55734, 140467, 25988, 4150, 7760, 27153	1.421552913	1	0.999997577	99.87966092
UP_SEQ_FEATURE	zinc finger region:C2H2-type 7	6	4.225352113	0.4020115	57473, 55734, 140467, 25988, 23394, 148213	1.433935778	1	0.999998596	99.95206845
INTERPRO	IPR015880:Zinc finger, C2H2-like	9	6.338028169	0.423624998	57473, 55734, 140467, 25988, 23394, 4150, 148213, 7760, 27153	1.250079189	1	0.999999737	99.94300263
SMART	SM00355:SM00355	9	6.338028169	0.434702026	57473, 55734, 140467, 25988, 23394, 4150, 148213, 7760, 27153	1.232288037	1	0.999965247	99.79812178
UP_KEYWORDS	Zinc	23	16.1971831	0.455817001	811, 57473, 8882, 55734, 10235, 140467, 91603, 5916, 54467, 25988, 27154, 7760, 27153, 51147, 57037, 8878, 64219, 4150, 23394, 84872, 7555, 148213, 221037	1.079873244	1	0.990365438	99.95303114
INTERPRO	IPR013087:Zinc finger C2H2-type/integrase DNA-binding domain	7	4.929577465	0.509234744	55734, 140467, 25988, 4150, 148213, 7760, 27153	1.218539103	1	0.999999842	99.99355617
UP_SEQ_FEATURE	zinc finger region:C2H2-type 4	6	4.225352113	0.572290618	55734, 140467, 23394, 4150, 7760, 27153	1.181736018	1	0.999999982	99.99967102
UP_SEQ_FEATURE	zinc finger region:C2H2-type 9	4	2.816901408	0.67089341	55734, 140467, 25988, 148213	1.161310951	1	1	99.99999331
UP_SEQ_FEATURE	zinc finger region:C2H2-type 8	4	2.816901408	0.734623539	55734, 140467, 25988, 148213	1.052194486	1	1	99.99999973
INTERPRO	IPR001909:Krueppel-associated box	3	2.112676056	0.751723232	148213, 7760, 27153	1.117672048	1	1	99.99999937
UP_KEYWORDS	Metal-binding	29	20.42253521	0.753684763	132, 8882, 5916, 25988, 5372, 6709, 57037, 8878, 64219, 4150, 9478, 51440, 23394, 148213,	0.932892837	1	0.996115194	99.99999783

					221037, 811, 57473, 55734, 140467, 10235, 47, 91603, 54467, 7760, 27154, 27153, 51147, 84872, 7555				
SMART	SM00349:SM00349	3	2.112676056	0.758522432	148213, 7760, 27153	1.101765317	1	0.999999999	99.99998066
GOTERM_MF_DIRECT	GO:0046872~metal ion binding	15	10.56338028	0.774268319	57473, 132, 55734, 140467, 47, 91603, 25988, 5372, 7760, 27153, 57037, 4150, 84872, 148213, 221037	0.90988025	1	0.99999782	99.999999622
UP_SEQ_FEATURE	zinc finger region:C2H2-type 6	4	2.816901408	0.829858128	55734, 140467, 25988, 27153	0.890778286	1	1	100
Annotation Cluster 11	**Enrichment Score: 0.2729863557501724**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	DNA repair	5	3.521126761	0.476976419	56949, 55031, 10273, 7520, 9557	1.398540742	1	0.991505834	99.971497422
GOTERM_BP_DIRECT	GO:0006281~DNA repair	6	4.225352113	0.528392009	56949, 91603, 10273, 7520, 25988, 9557	1.240100251	1	0.999999938	99.999190322
UP_KEYWORDS	DNA damage	5	3.521126761	0.601987929	56949, 55031, 10273, 7520, 9557	1.195116634	1	0.993223328	99.999086221
Annotation Cluster 12	**Enrichment Score: 0.26576828453800444**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_CC_DIRECT	GO:0000139~Golgi membrane	7	4.929577465	0.371883966	55697, 22883, 10565, 64689, 9478, 3382, 1174	1.408271144	1	0.999491259	99.721708221
GOTERM_CC_DIRECT	GO:0030054~cell junction	5	3.521126761	0.459244565	10235, 22883, 9478, 3382, 9223	1.430624655	1	0.999831573	99.95816593
UP_KEYWORDS	Synapse	4	2.816901408	0.483045592	10235, 22883, 9478, 3382	1.531605783	1	0.990784417	99.975393936
UP_KEYWORDS	Cell junction	6	4.225352113	0.660219967	10235, 23224, 22883, 9478, 3382, 9223	1.07073798	1	0.992046071	99.999875337
UP_KEYWORDS	Golgi apparatus	6	4.225352113	0.861053867	22883, 10565, 64689, 9478, 3382, 1174	0.818799632	1	0.998983221	100
Annotation Cluster 13	**Enrichment Score: 0.2241334502934169**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**

GOTERM_BP_DIRECT	GO:0018279~protein N-linked glycosylation via asparagine	4	2.816901408	0.419016542	811, 64689, 5372, 6709	1.689025431	1	0.999999472	99.97904552
GOTERM_BP_DIRECT	GO:0043687~post-translational protein modification	5	3.521126761	0.686323375	811, 1938, 64689, 5372, 6709	1.072946483	1	0.999999993	99.9999986
GOTERM_BP_DIRECT	GO:0044267~cellular protein metabolic process	9	6.338028169	0.739332317	811, 9804, 8669, 1938, 1975, 64689, 5372, 6175, 6709	0.94499618	1	0.999999998	99.99999992
Annotation Cluster 14	**Enrichment Score: 0.22004753852343656**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_BP_DIRECT	GO:0015031~protein transport	5	3.521126761	0.544694798	25909, 51271, 10565, 64689, 51510	1.283328146	1	0.999999936	99.99953228
UP_KEYWORDS	Transport	14	9.85915493	0.585915927	782, 9804, 3382, 25909, 23307, 65110, 10650, 51271, 10565, 6447, 64689, 51510, 7385, 1174	1.039818992	1	0.993559858	99.99849555
UP_KEYWORDS	Protein transport	7	4.929577465	0.685281257	25909, 9804, 51271, 10565, 64689, 51510, 1174	1.018715655	1	0.992997531	99.99995252
Annotation Cluster 15	**Enrichment Score: 0.21010624314667864**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	mRNA splicing	5	3.521126761	0.515127509	10421, 5511, 56949, 8882, 10147	1.332393545	1	0.991854001	99.98902005
GOTERM_BP_DIRECT	GO:0006397~mRNA processing	3	2.112676056	0.632090492	5511, 8882, 10147	1.402494331	1	0.999999962	99.99998317
UP_KEYWORDS	mRNA processing	5	3.521126761	0.639522071	10421, 5511, 56949, 8882, 10147	1.139851125	1	0.991578556	99.99973754
GOTERM_BP_DIRECT	GO:0008380~RNA splicing	4	2.816901408	0.693466506	10421, 5511, 8882, 10147	1.121995465	1	0.999999993	99.99999902
Annotation Cluster 16	**Enrichment Score: 0.1922732305315986**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	S-adenosyl-L-methionine	3	2.112676056	0.514553303	154743, 29072, 4839	1.739949217	1	0.99300418	99.98885515
UP_KEYWORDS	Methyltransferase	3	2.112676056	0.522415941	154743, 29072, 4839	1.714732562	1	0.991412174	99.99092688
UP_KEYWORDS	Transferase	8	5.633802817	0.985674493	132, 65268, 154743, 47, 29072, 6817, 3932, 4839	0.588639536	1	0.999995532	100

Annotation Cluster 17	**Enrichment Score: 0.1785082046228081**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_MF_DIRECT	GO:0000166~nucleotide binding	5	3.521126761	0.557636129	9987, 1975, 65110, 10137, 9557	1.263057453	1	0.999918649	99.99756304
UP_KEYWORDS	RNA-binding	9	6.338028169	0.571834908	5511, 170506, 9987, 1975, 65110, 26135, 10137, 4150, 4839	1.098915295	1	0.993211574	99.99770762
INTERPRO	IPR012677:Nucleotide-binding, alpha-beta plait	4	2.816901408	0.63400978	9987, 1975, 65110, 10137	1.226573427	1	0.999999997	99.99987924
INTERPRO	IPR000504:RNA recognition motif domain	3	2.112676056	0.794603548	9987, 1975, 10137	1.022144522	1	1	99.99999995
SMART	SM00360:SM00360	3	2.112676056	0.797243436	9987, 1975, 10137	1.016283525	1	0.999999999	99.99999711
Annotation Cluster 18	**Enrichment Score: 0.16844033939294323**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	Mitosis	4	2.816901408	0.545676297	84722, 993, 157313, 91603	1.396065449	1	0.992736776	99.99516218
UP_KEYWORDS	Cell division	5	3.521126761	0.625715889	84722, 25909, 993, 157313, 91603	1.159966145	1	0.99193407	99.99957866
GOTERM_BP_DIRECT	GO:0051301~cell division	4	2.816901408	0.70723602	84722, 993, 157313, 91603	1.098457098	1	0.999999993	99.99999952
UP_KEYWORDS	Cell cycle	7	4.929577465	0.74667333	84722, 23741, 25909, 993, 51147, 157313, 91603	0.951972215	1	0.996097539	99.99999691
GOTERM_BP_DIRECT	GO:0007067~mitotic nuclear division	3	2.112676056	0.797610788	993, 157313, 91603	1.015599343	1	1	100
Annotation Cluster 19	**Enrichment Score: 0.08782667190035684**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	8	5.633802817	0.354247805	811, 182, 10235, 22883, 1954, 9478, 51440, 6709	1.377468935	1	0.997911576	99.66394295
INTERPRO	IPR013320:Concanavalin A-like lectin/glucanase, subgroup	3	2.112676056	0.500714769	811, 22883, 23231	1.784938942	1	0.999999922	99.99186266
GOTERM_CC_DIRECT	GO:0005576~extracellular region	7	4.929577465	0.619850884	811, 6352, 182, 22927, 6696, 6447, 27154	1.090274434	1	0.99986564	99.99951588
UP_KEYWORDS	Secreted	4	2.816901408	0.987066664	811, 6352, 6696, 6447	0.507252076	1	0.999996043	100

UP_SEQ_FEATURE	disulfide bond	5	3.521126761	0.997884671	811, 6352, 182, 6447, 1954	0.427884766	1	1	100
UP_SEQ_FEATURE	signal peptide	7	4.929577465	0.999429654	811, 6352, 182, 6696, 22883, 6447, 1954	0.425363895	1	1	100
UP_KEYWORDS	Signal	7	4.929577465	0.999573434	811, 6352, 182, 6696, 22883, 6447, 1954	0.416398103	1	0.999999999	100
UP_KEYWORDS	Disulfide bond	5	3.521126761	0.999740886	811, 6352, 182, 6447, 1954	0.354665908	1	1	100
UP_SEQ_FEATURE	topological domain:Extracellular	3	2.112676056	0.999760153	182, 22883, 1954	0.288901066	1	1	100
UP_KEYWORDS	Glycoprotein	8	5.633802817	0.999899872	811, 6352, 5688, 182, 6696, 22883, 1954, 23231	0.39786985	1	1	100
UP_SEQ_FEATURE	glycosylation site:N-linked (GlcNAc...)	6	4.225352113	0.99997295	811, 182, 6696, 22883, 1954, 23231	0.327528506	1	1	100
Annotation Cluster 20	**Enrichment Score: 0.040079534651219315**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
GOTERM_CC_DIRECT	GO:0005743~mitochondrial inner membrane	3	2.112676056	0.853299126	3030, 30968, 7385	0.891389208	1	0.999996997	100
UP_KEYWORDS	Mitochondrion	6	4.225352113	0.938679623	9804, 3030, 23224, 10650, 30968, 7385	0.69190963	1	0.999856239	100
UP_KEYWORDS	Transit peptide	3	2.112676056	0.946548136	3030, 30968, 7385	0.653682579	1	0.999892922	100
Annotation Cluster 21	**Enrichment Score: 0.013698958690022247**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**
UP_KEYWORDS	ATP-binding	10	7.042253521	0.952002596	132, 1615, 65268, 170506, 65985, 47, 7520, 9557, 3932, 9223	0.704265159	1	0.999884451	100
UP_KEYWORDS	Nucleotide-binding	12	8.450704225	0.96934964	132, 1615, 65268, 170506, 65985, 47, 1938, 7520, 9557, 26354, 3932, 9223	0.685893025	1	0.99996486	100
GOTERM_MF_DIRECT	GO:0005524~ATP binding	10	7.042253521	0.976434739	132, 1615, 65268, 170506, 65985, 47, 7520, 9557, 3932, 9223	0.64859707	1	1	100
UP_SEQ_FEATURE	nucleotide phosphate-binding region:ATP	6	4.225352113	0.978233108	65268, 170506, 47, 9557, 3932, 9223	0.587913669	1	1	100
Annotation Cluster 22	**Enrichment Score: 4.4889251273977734E-4**								
Category	**Term**	**Count**	**%**	**P-Value**	**Genes**	**Fold Enrichment**	**Bonferroni**	**Benjamini**	**FDR**

GOTERM_CC_DIRECT	GO:0016021~integral component of membrane	8	5.633802817	0.996073634	23224, 22883, 5916, 1954, 162427, 23231, 57523, 113201	0.518481418	1	1	100
UP_SEQ_FEATURE	topological domain:Cytoplasmic	6	4.225352113	0.999431496	182, 9804, 23224, 22883, 1954, 113201	0.4033713	1	1	100
UP_SEQ_FEATURE	transmembrane region	10	7.042253521	0.999543607	1831, 182, 9804, 23224, 22883, 1954, 162427, 23231, 57523, 113201	0.47165156	1	1	100
UP_KEYWORDS	Transmembrane helix	9	6.338028169	0.999890144	182, 9804, 23224, 22883, 1954, 162427, 23231, 57523, 113201	0.418079671	1	1	100
UP_KEYWORDS	Transmembrane	9	6.338028169	0.999901058	182, 9804, 23224, 22883, 1954, 162427, 23231, 57523, 113201	0.415631897	1	1	100

Appendix E – ImageJ quantitation of scFv concentration

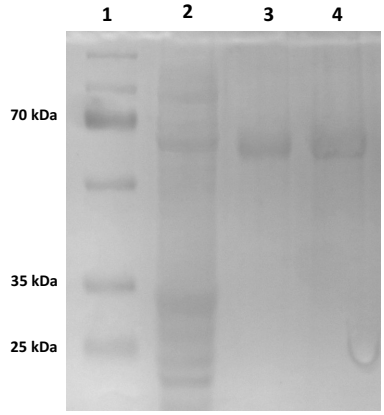

Figure 8.8 –Comparison of P3P2F10 scFv to BSA standards on an SDS-PAGE gel for determination of scFv concentration using ImageJ software. Lane 1 contains protein ladder to allow for size determination of fragments. Lane 2 contains 15μL of the purified P3P2F10 scFv. Lane 3 contains 15μL of 1 μg/mL BSA. Lane 4 contains 15 μL of 2 μg/mL BSA. The image was converted to greyscale during quantitation.