

Development and validation of an autoantibody-based blood test for hepatocellular carcinoma

Christopher James Welberry, MSci Hons

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, September 2018

Contents

Abstract.....I

Acknowledgments II

List of figures.....III

List of tables..... VIII

Abbreviations.....XI

Chapter 1. Introduction 1

1.1 Liver biology 2

1.1.1 Structure 2

1.1.2 Function 3

1.2 Hepatocellular Carcinoma.....5

1.2.1 HCC epidemiology.....6

1.2.2 High risk groups 8

1.2.3 Molecular pathogenesis 17

1.2.4 Screening 19

1.2.5 Diagnosis.....23

1.2.6 Staging 25

1.2.7 Treatments 26

1.3 Immune system.....27

1.4 Autoimmunity and Autoantibodies 31

1.5 Autoantibodies in cancer 32

1.5.1 Mutated TAAs.....32

1.5.2 Aberrantly expressed TAA 33

1.5.3 Post-translationally modified TAAs. 33

1.5.4 Autoantibody investigation methods.....33

1.5.5 Clinical utility of Autoantibodies as biomarkers.....34

1.6 Autoantibodies in HCC 35
Chapter 2. Aims of the project 36
Chapter 3. High-throughput antigen production 38
3.1 Introduction 38
3.2 Methods 38
3.2.1 Identification of potential biomarkers from the literature 38
3.2.2 Antigen Cloning 40
3.2.3 High-throughput protein expression 45
3.2.4 Full scale expression 46
3.2.5 HTP Purification 46
3.2.6 Full scale purification 48
3.3 Results 48
3.3.1 Antigen selection 48
3.3.2 Cloning 60
3.3.3 Expression and Purification 67
3.4 Discussion 78
Chapter 4. Discovery assay 81
4.1 Introduction 81
4.2 Methods 81
4.2.1 Serum samples 81
4.2.2 2µM stock plate preparation 82
4.2.3 Plate Coating 83
4.2.4 High Throughput Assay protocol 85
4.2.5 Verification of plate layout 85
4.2.6 AFP measurement 86
4.2.7 Data analysis 86

4.3 Results 87
4.3.1 Stock plate confirmation 87
4.3.2 High Throughput Assay results 91
4.4 a 96
Chapter 5. Lead antigen reproduction 99
5.1 Introduction 99
5.2 Methods 101
5.2.1 Expression 101
5.2.2 Purification 102
5.3 Results 104
5.3.1 Expression 104
5.3.2 Purification 105
5.4 Discussion 107
Chapter 6. Combining leads 108
6.1 Introduction 108
6.2 Methods 109
6.2.1 Serum samples 109
6.2.2 Plate coating 110
6.2.3 Assay protocol 111
6.2.4 Data analysis 111
6.3 Results 112
6.3.1 Individual markers 112
6.3.2 Net reclassification improvement analysis 115
6.3.3 Addition of AFP 116
6.3.4 Performance by stage 116
6.3.5 Assay reproducibility 117

6.4 Discussion....120
Chapter 7. Training and validation....122
7.1 Introduction....122
7.2 Methods....122
7.2.1 Serum samples....122
7.2.2 Plate coating....123
7.2.3 Assay protocol....123
7.2.4 AFP measurement....123
7.2.5 Data analysis....124
7.3 Results....125
7.3.1 Training....125
7.3.2 Validation....127
7.4 Discussion....130
Chapter 8. Preliminary serial samples assay....133
8.1 Intro....133
8.2 Methods....133
8.2.1 Antigen selection....133
8.2.2 Autoantibody ELISA....133
8.3 Results....133
8.4 Discussion....136
Chapter 9. Discussion....138
Chapter 10. References....147
Chapter 11. Appendices....165
11.1 Medias....165
11.2 Purification buffers....165
11.3 Stock solutions....167

11.4 Assay buffers 167

11.5 Copyright permissions 168

Abstract

Hepatocellular carcinoma is the 4th leading cause number of cancer related deaths worldwide. Detection of Hepatocellular carcinoma (HCC) at an early stage is crucial for a positive prognostic outcome, however early detection is hindered by the paucity of early symptoms and interference of underlying liver disease with imaging modalities. Assessment of Alpha-fetoprotein (AFP) as a blood-based screening/diagnostic tool for HCC is still under scrutiny as many HCC tumours do not produce elevated levels and levels can be elevated in other liver disease patients.

Autoantibodies to tumour associated antigens have been shown to be present in the blood of patients with numerous solid malignancies including HCC and, in some cases, have been reported to be detectable before clinical diagnosis. This thesis will look to expand on previously published work which demonstrates the presence and utility of such autoantibodies to detect HCC from high-risk liver disease patients and how this could fit into the clinical pathway for such patients.

As such, this thesis documents the process of characterising autoantibody responses in such patients to previously described tumour-associated antigens along with the characterisation of novel autoantigens. Autoantibody responses were measured using ELISA against recombinant antigens produced in *E. coli*. The final autoantibody panel was formed through multiple studies including training and validation of the final panel in independent patient cohorts.

A final panel of 11 autoantibodies was selected from training assays and validated in a separate cohort with no significant difference in accuracy. The panel of 11 could detect HCC with a sensitivity of 42% and a specificity of 88.2% in the training and validation cohorts combined. AFP measurement was significantly additive to the 11-autoantibody panel resulting in sensitivity and specificity of 56.3% and 87.8% respectively. Importantly the combined panel sensitivity for BCLC stage A HCCs, was greater than 55%, with current tests often showing stage A sensitivities of less than 50%.

This thesis shows autoantibodies can distinguish HCC from patients with chronic liver and most importantly at stages with a favourable prognosis. Autoantibodies could also be combined with AFP antigen measurements to improve on either test alone. This thesis also brings this autoantibody panel closer to a commercial test by assessing performance in independent training and validation cohorts.

Acknowledgments

I would like to thank my supervisors, Dr Caroline Chapman and Professor Will Irving and to the University of Nottingham for this chance to be involved in such an exciting step in my scientific career. Huge thanks to Will for his critical review of this manuscript.

I am especially grateful to everyone at Oncimmune, firstly for the funding but most importantly for facilitating this project and allowing it to run smoothly from start to finish. I would sincerely like to thank Dr. Andrea Murray for the constant faith she has shown in me over the past four years. Thanks to Drs. Isabel Macdonald and Celine Parsy-Kowalska for all their significant scientific help and guidance. Thanks to Jane McElveen and everyone that has helped to facilitate the running of the assays involved. Thanks to Jared Allen, Graham Healey and Chris Reynolds for all the help compiling and understanding the data.

Thanks to my undergraduate project supervisor Professor Conrad Nieduszynski, thanks for inspiring me to follow a career in science and for the all the effort you put in to gaining me invaluable research experience that prepared me for this experience.

I couldn't have made it through this without out all the friends surrounding me that helped take my mind off all the work when needed, I can't name you all but I'm sincerely grateful to have you all.

Thanks to all the members of the University of Nottingham frisbee team, I will always cherish our national championship!

Thanks to all my friends past and present at Oncimmune, especially Jen, Leanna, Jared, Adam, Chris, Simon, Hayley, Stacy, Nat, Tom, Leon and Jacek. Thanks for all the laughs and the board games.

Most of I would like to thank and dedicate this thesis to my family, thanks for being there for me through the last 8 years and always believing in me!

List of figures

Figure 1 - Distribution of Cases and Deaths for the 10 Most Common Cancers (excluding nonmelanoma skin cancers) in 2018 for Males and Females[1]. 1

Figure 2: Diagram of a Hepatic lobule modified from Gulec et al. (2010)[4].....3

Figure 3 - Chronological representation of histopathological events during progression toward HCC. The majority of HCC lesions occur against a cirrhotic background (green asterisk), whereas in rare instances it can also occur against a noncirrhotic background (red asterisk). Typical histopathology sections during disease progression are shown: (a) normal liver (×200), (b) chronic hepatitis with marked portal/peri-portal chronic inflammation (×100), (c) cirrhosis with regenerating parenchymal nodules separated by active fibrous septa (×40), (d1) ductular reaction in cirrhosis (×100), (d2) ductular reaction with different stages of differentiation toward hepatocytes in a regenerating nodule (×200), (d3) large cell changes in regenerating nodule in cirrhosis (×100), (e1) large cell dysplasia (×200), (e2) small cell dysplasia (×200), (f1) HCC with thick trabeculae of malignant hepatocytes (×200) and (f2) HCC with trabeculae and pseudoacinar transformation in a noncirrhotic background (×40). Reproduced with permission from Ramakrishna et al. (2013)[22].....9

Figure 4 - HBV genome based on the reference sequence NC_003977.2 14

Figure 5 - 54 year old male with hepatitis C cirrhosis. CT shows an arterial enhancing nodule a with washout of contrast in the delayed phase b consistent with hepatocellular carcinoma. Willatt et al. (2017) [97] 24

Figure 6 - BCLC staging system. Modified from Forner, Reig & Bruix (2018)[97] 26

Figure 7 – Expression region of the NLIC and CLIC vectors used. Antigen insertion relates to the linearization site of the vector, where the antigen DNA is inserted 41

Figure 8 - Selection of appropriate AAb papers 49

Figure 9 - Insert amplification results from the first attempt with standard PCR protocol. Failed reactions are highlighted with a red box 61

Figure 10- Insert amplification results from the first attempt with standard PCR protocol. Failed reactions are highlighted with a red box 62

Figure 11 - TDPCR insert amplification results. Failed reactions are highlighted with a red box 63

Figure 12 – Standard PCR amplification of remaining inserts using reprepared template DNA 64

Figure 13 – Example colony PCR gel. Labels contain antigen name followed by expected size in BPs.....66

Figure 14 – SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. 69

Figure 15 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification 70

Figure 16 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project. 71

Figure 17 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project. 72

Figure 18 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project. 73

Figure 19- SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification 74

Figure 20- SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project. 75

Figure 21- Anti-HIS concentration ELISA for the five batches which showed concentration of 50-100µg/ml by Bradford assay 77

Figure 22 - Layout of the 2µM, 160nM and 50nM stock plates 83

Figure 23 - Layout of the 96 well plates. One 96-well plate was used per sample 84

Figure 24 – Legend for figure 25, showing the confirmation Coomassie stained gels and anti-HIS westerns of the stock plate layout. 89

Figure 25 –Coomassie stained gels (top) and anti-HIS westerns (bottom) of the stock plate layout. Red boxes highlight the band at the expected size 90

Figure 26 – Dot plot results of a selection of AAbs run in this study split by sample type 92

Figure 27 - Example chromatogram from an AKTA prime plus. Blue line shows the UV measurement of the liquid leaving the column and is representative of the protein content of the liquid 100

Figure 28 – Dot plot of AAb reactivity of three groups of samples, ovarian cancer, high risk controls and healthy controls against two batches of antigen. One batch (HTP) was produced using the HTP technique described in chapter 4 and the full-scale batch produced using the technique described in section 5. Samples highlighted in red were positive with either batch of antigen. 100

Figure 29 – Anti-HIS western blots of each protein. Lane 1 of each blot contains Novex sharp prestained protein standard. Red boxes highlight bands of correct MW. Lane D2 contains a protein purified for a separate Oncimmune project. 106

Figure 30 - Assay plate one layout.....111

Figure 31 - Assay plate two layout 111

Figure 32 - Assay plate three layout 111

Figure 33 - Dot plot results of a selection of AAbs run in this study split by sample type 115

Figure 34 – Example Bland-Altman plots showing 3 situations of antigen reproducibility between this assay and the discovery assay. Plot A shows NY-ESO-1 at 50nM, for this, the same batch of antigen was used in each assay, confidence intervals are tight and samples which appear to be true positives (red circle) results lie close to a difference in signal of 0. Plot B shows HSPA4 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however samples which appear to be true positives (red circle) have a high mean signal and difference in signal between the two batches close to 0. Plot C shows NPM1 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however all signal appeared to be from non-specific binding as there are no samples with a high mean signal and difference between the two batches close to 0. 120

Figure 35 - OD readings at a protein concentration on 160nm of multiple samples from Patient 1. Patient 1 is male and 46 years old at the date of first sample collection. HCC was clinically after the date of final sample collection. 134

Figure 36 - OD readings at a protein concentration on 160nm of multiple samples from Patient 2. Patient 2 is female and 63 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken. 135

Figure 37 - OD readings at a protein concentration on 160nm of multiple samples from Patient 3. Patient 3 is female and 43 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken. 135

Figure 38 - OD readings at a protein concentration on 160nm of multiple samples from Patient 4. Patient 4 is male and 68 years old at the date of first sample collection. HCC was clinically diagnosed after the date of final sample collection. 136

Figure 39 - OD readings at a protein concentration on 160nm of multiple samples from Patient 5. Patient 5 is female and 52 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken. 136

Figure 40 - Theoretical use of the AAb+AFP panel in a cohort of 100 high risk patients at a risk of HCC of 4%. Modelled increases in risk are based on the sensitivity and specificity of ultrasound and the AAb+AFP panel shown in the figure.....141

Figure 41 - Theoretical use of the AAb+AFP panel in a cohort of 100 patients with a liver lesion identified by Ultrasound screening at a HCC risk of 23%. Modelled increases in risk are based on the sensitivity and specificity of ultrasound, CT, MRI and the AAb+AFP panel shown in the figure.....142

List of tables

Table 1 – Comparison of Area under the receiver operating characteristic curve (AUC) of screening tests for HCC for all stages of tumour. 20

Table 2 - Score assigned for each factor in the literature review. Final score for each potential marker was the sum of each factor.....39

Table 3 - Primer design 43

Table 4 – PCR reagents for insert amplification 43

Table 5 – Standard PCR conditions.....43

Table 6 – Touchdown PCR conditions 43

Table 7 - Heat shock protocol.....44

Table 8 - Colony PCR reagents.....44

Table 9 - Colony PCR conditions 45

Table 10 - BL21(DE3) RIPL transformation protocol.....45

Table 11 - Selected antigens and their link to HCC. For the rest of this thesis all antigens will be referred to by their short protein name. 55

Table 12 – Protein concentration of each batch of protein chosen for discovery assay.....76

Table 13 - Age and gender data for discovery assay samples.....82

Table 14 - HCC stage of discovery assay samples. N/A = stage data unavailable 82

Table 15 - Aetiology of the NCCLD controls 82

Table 16 – Number of positive samples in each sample group for each AAb..93

Table 17 – NRI formation of the AAb panel, sensitivity and specificity values shown are cumulative as each marker is added to the panel.....95

Table 18 – Performance of the AAb panel, AFP and the combination of both by HCC stage.....96

Table 19 – Pellet weight and lysis method chosen for each antigen.....105

Table 20 – Concentration and yields of the antigen batches produced.....106

Table 21 - Age and gender data of the assay samples by group.....109

Table 22 - Primary tumour size of the HCC patients (where available).....109

Table 23 - HCC stage of the assay samples. N/A = stage data unavailable....109

Table 24 - Aetiology of the NCCLD controls.....110

Table 25 – Number of positives and calculated sensitivity and specificity of each marker by sample group. N/A implies that no cut-off could be assigned which matched the criteria.....113

Table 26 - NRI formation of the AAb panel, sensitivity and specificity values shown are cumulative as each marker is added to the panel.....116

Table 27 - Performance of the AAb panel, AFP and a combination of both by HCC TNM stage.....117

Table 28 - Parameters established from Bland-Altman plots for each antigen/concentration ranked by the 95% confidence interval range.....119

Table 29 - Training set sample origins.....123

Table 30 - Validation set sample origins.....123

Table 31 - Performance of the individual AAbs at the cut-offs selected by the Monte Carlo direct search in the training set.....126

Table 32 - Performance of the AAb panel by BCLC stage in the training set .126

Table 33 - Performance of AFP (>200ng/ml) by BCLC stage in the training set 127

Table 34- Performance of the AAb panel in combination with AFP by BCLC stage in the training set.....127

Table 35 - Performance of the individual AAbs in the validation set at the cut-offs selected in the training set 128

Table 36 - Performance of the AAb panel by BCLC stage in the validation set 129

Table 37 - Performance of AFP by BCLC stage in the validation set.....129

Table 38 - Performance of the AAb panel in combination with AFP by BCLC stage in the training set.....130

Table 39 – Sensitivity of the AAb panel, AFP and a combination of both by BCLC stage in the combined training and validation sets**Error! Bookmark not defined.**

Abbreviations

1M dialysis buffer (1MDB)

2M Dialysis buffer (2MDB)

3, 3´, 5, 5´-tetramethylbenzidine (TMB)

4M dialysis buffer (4MDB)

Aflatoxin B1 (AFB1)

Alpha-fetoprotein (AFP)

American Association for the Study of Liver Diseases (AASLD)

Amino acid (AA)

Autoantibody (AAb)

Autoimmune hepatitis (AIH)

Barcelona-Clinic liver cancer (BCLC)

Base-pair (BP)

B cell receptor (BCR)

Carbenicillin (CB)

Chloramphenicol (CM)

Computed tomography (CT)

Des-gamma-carboxy prothrombin (DCP)

Diethylnitrosamine (DEN)

Distilled water (dH2O)

Ductal reactions (DRs)

E. coli BirA biotin-protein ligase enzyme recognition sequence (BirA)

Enzyme-linked immunosorbent assay (ELISA)

European Association for the Study of the Liver (EASL)

European Organisation for Research and Treatment of Cancer (EORTC)

Extracellular matrix (ECM)

Fast protein liquid chromatography (FPLC)

Glypican-3 (GPC3)

Golgi protein 73 (GP73)

Guanidine binding buffer (GBB)

Hepatitis B e antigen (HBeAg)

Hepatitis B surface antigen (HBsAg)

Hepatitis B virus (HBV)

Hepatitis B x antigen (HBxAg)

Hepatitis C virus (HCV)

Hepatocellular carcinoma (HCC)

Hereditary hemochromatosis (HH)

Hormone replacement therapy (HRT)

Long-lived plasma cells (LLPCs)

Magnetic resonance imaging (MRI)

Major histocompatibility complex class 1 (MHC1)

Micro-RNA (miRNA)

Model for end stage liver disease (MELD)

Natural killer (NK)

Net reclassification improvement (NRI)

Nicotinamide adenine dinucleotide (NAD)

Non-alcoholic fatty liver disease (NAFLD)

Non-alcoholic Steatohepatitis (NASH)

Non-cancerous chronic liver disease (NCCLD)

Open reading frame (ORF)

Optical density (OD)

Osteopontin (OPN)

Phosphate buffered Turbo Broth Overnight expression medium (TBOnex-P)

Prolactin (PRL)

Quality adjusted life year gained (QALY)

Randomised control trial (RCT)

Reactive oxygen species (ROS)

Regenerating nodules (RNs),

Relative units (RU)

Room temperature (RT)

Serological analysis of expression cDNA libraries (SEREX)

Serological proteomics-based approach, SERPA

Streptomycin (STREP)

Touchdown PCR (TDPCR)

Transcatheter arterial chemoembolization (TACE

Tumour-associated antigen (TAA)

Ultrasound (US)

United network for organ sharing (UNOS)

Urea binding buffer (UBB)

Urea column scrubbing buffer (USCB)

Urea dialysis buffer 2 (UDB2)

Urea elution buffer (UEB)

Vector only lysate (VOL)

Chapter 1. Introduction

Hepatocellular carcinoma is one of the major challenges of modern oncology. It is the fifth most common cancer in men and 9th in women yet ranks 2nd in the most common causes of cancer related deaths worldwide for men and 6th for women (Figure 1)[1]. The majority of cases occur in countries with high prevalence of Hepatitis B or C viral infection, such as China, Japan and Egypt. However; rates in more developed western countries are on the rise, attributed to modern lifestyle changes such as increased alcohol consumption and poor diets.

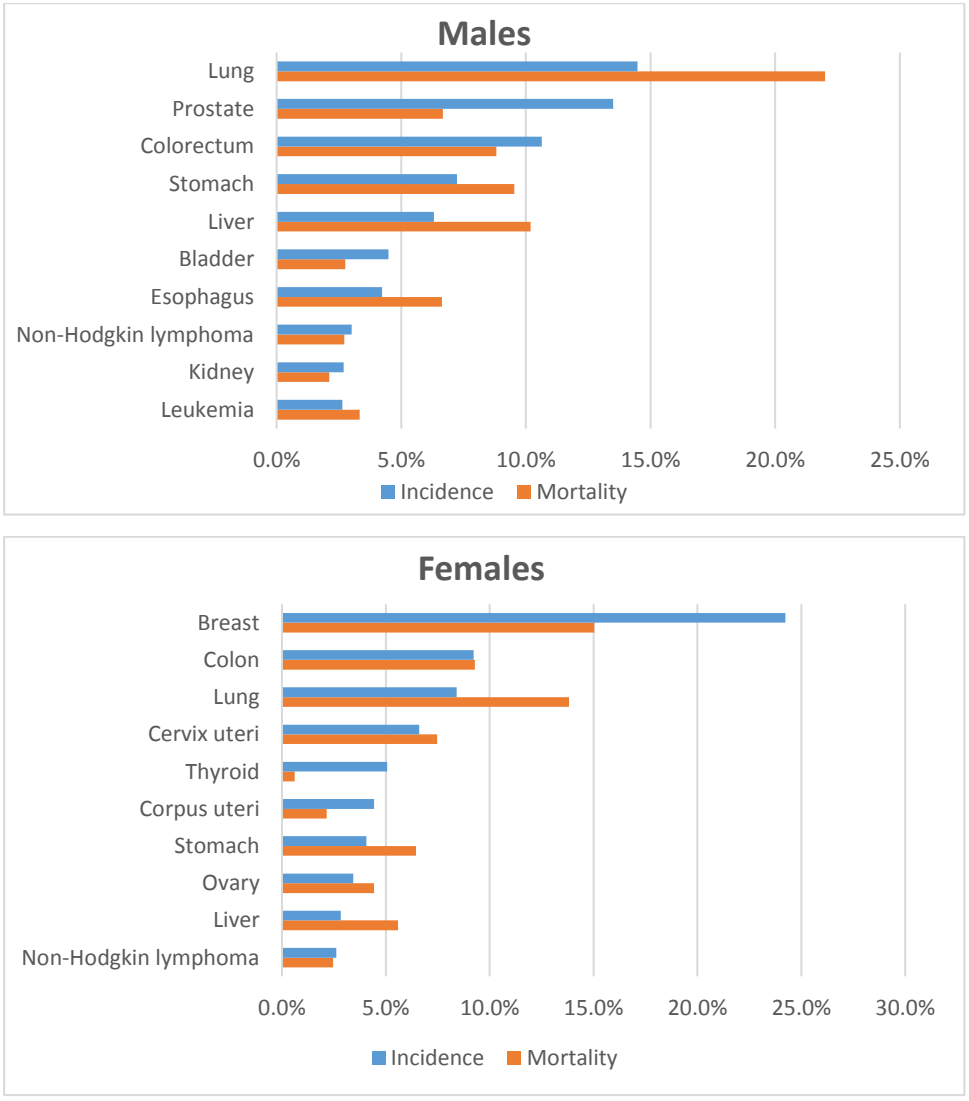


Figure 1 - Distribution of Cases and Deaths for the 10 Most Common Cancers (excluding nonmelanoma skin cancers) in 2018 for Males and Females[1].

1.1 Liver biology

1.1.1 Structure

The liver, lying in the upper right section of the abdomen, is the largest internal organ in the human body, normally weighing 968-1860g in men[2] and 603 to 1767 g in women[3]. It holds, on average, 12 % of the body's blood supply, leading to a deep red/brown appearance. The liver is composed of two major lobes separated by the falciform ligament. Oxygenated blood is supplied to the liver via the hepatic artery. The liver also receives deoxygenated blood, carrying nutrients and toxins to be processed, through the portal vein from the stomach and small intestines. All blood leaves the liver via the hepatic vein into the inferior vena cava. Approximately 25% of the blood supplied is oxygenated, meaning the majority is unoxygenated blood delivered for processing. Each lobe is made up of multiple lobules. Lobules are hexagonal structures with a central vein (Figure 2). The lobules are primarily made up of hepatocytes, the cells responsible for the primary functions of the liver. Discontinuous capillaries, known as sinusoids, allow blood to enter the space of Disse, percolate around the hepatocytes, then facilitate removal after processing. At each corner of the hexagon lies a portal triad. Portal triads each contain a branch of the hepatic artery, portal vein and the bile duct. Portal triads are responsible for carrying blood to and from the lobules, as well as transporting bile produced in the liver to the gall bladder.

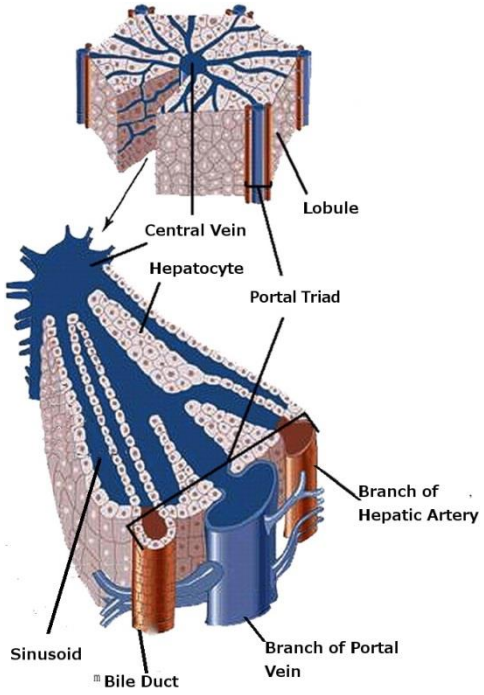


Figure 2: Diagram of a Hepatic lobule modified from Gulec et al. (2010)[4].

1.1.2 Function

1.1.2.1 *Synthesis of amino acids*

The liver is responsible for synthesis and breakdown of many key amino acids. If protein is in excess, the liver deaminates essential amino acids into alpha-keto acids via oxidative deamination, removing the nitrogen containing amino group by oxidation. Glutamic acid derived from the amination of alpha-ketoglutarate by ammonia, can be used in the amination of other alpha-keto acids to form corresponding non-essential amino acids[5]. This is a key step in utilising nitrogen as ammonia is toxic, so most of the ammonia not used in this process is converted to urea for excretion in urine. Alpha-keto acids can also be converted into lipids, glucose or glycogen.

1.1.2.2 *Detoxification*

The liver has two phases of detoxification for harmful substances. Phase one detoxification is catalysed by the cytochrome P450 enzyme group, a group of membrane bound oxidase enzymes[6]. The liver aims to oxidise toxins into a less harmful form; however, the by-product of this is the production of free radicals which can be damaging to the liver cells. The oxidation process can make some

toxins more dangerous, for example aflatoxin, which when oxidised forms a very potent carcinogen[7].

Phase two detoxification relies on conjugation of another substance such as cysteine, glycine or sulphur. This process should render the toxin or drug less harmful as well as making it water soluble, so it can be excreted via bile or urine.

1.1.2.3 *Bile production*

Hepatocytes are responsible for the production of bile. Bile is made up of water, bile acids and bile pigments. The role of bile acids is to emulsify fats, facilitating easier breakdown and absorption in the small intestines. Bile pigment is a result of the breakdown of bilirubin. When blood cells die some components are recycled in the spleen, however during the breakdown of heme, separation of the iron leaves behind an unrecyclable molecule, unconjugated bilirubin. Unconjugated bilirubin is toxic and must be recycled by the liver[8]. Unconjugated bilirubin is lipid soluble so requires albumin to carry it to the liver for further metabolism. The liver conjugates the bilirubin to glucuronic acid producing conjugated bilirubin which is water soluble. This is then transported to the gall bladder to be secreted into the small intestine with the other components of bile.

1.1.2.4 *Synthesis of protein*

Hepatocytes are responsible for the production of most plasma proteins. Albumin is the major output of the liver at a rate of 9-12g/day. Albumin is needed to bind and transport bilirubin, as well as playing a major role in maintaining blood pressure. Albumin is also responsible for transporting other lipid soluble molecules such as hormones and drugs. Other plasma proteins are also produced in the liver such as transferrin (carries iron ions), IGFBP (carries IGF1) and clotting factors. The liver plays a role in the immune system by producing C-reactive protein and complement c1-c9[9], as well as the endocrine system producing hormones and prohormones including IGF[10], thrombopoeitin[11] and angiotensin[12].

1.1.2.5 *Carbohydrate and lipid metabolism*

The liver is crucial in maintaining correct blood glucose concentrations. It is a major site of glycogenesis when blood glucose levels are high, and can perform glycogenolysis when glucose levels are low[13]. It is also able to perform gluconeogenesis when the body is starved of glucose reserves. The liver is responsible for fructose breakdown, where it is converted into triglycerides and exported to other tissues to be metabolised. The liver is the main producer of triglycerides to be exported and stored in liver adipose tissue. The recent obesity trends have seen an increase in non-alcoholic fatty liver disease (NAFLD) where these triglycerides build up in hepatocytes causing liver damage and cirrhosis[14].

1.1.2.6 *Regeneration*

The liver has the highest potential for complete regeneration of any visceral organ[15]. Due to its role in toxin control, hepatocytes are prone to dying and therefore must be replaced. Whilst this is critical in maintaining the crucial function of the liver, it leaves the liver prone to cancer development. Hepatocytes undergo a very high turnover rate alongside exposure to many chemicals that are damaging to DNA. This leaves the hepatocytes prone to incorporating mutations, often resulting in carcinogenesis.

1.2 Hepatocellular Carcinoma

Hepatocellular carcinoma (HCC) is the primary malignancy of the liver, accounting for >90% of all liver cancer cases[16]. These rates are not surprising considering hepatocytes make up 70-80%[17] of the liver's mass. Cholangiocarcinoma, a cancer of mutated bile duct epithelial cells, is the second most common form of liver cancer yet only accounts for approximately 6% of all liver cancers. Other rare liver cancers include hepatoblastoma and more generalised malignancies such as blood vessel, muscle and connective tissue tumours. Basic classification of HCC separates tumours into early and progressed HCC. Early tumours are <2cm in size, of low grade/well differentiated (G1) and only vaguely nodular without clear vascularisation. Progressed HCCs are >2cm in size, distinctly nodular and show vascular

invasion. 5-year survival is much higher in early HCC compared with progressed HCC, 89% vs 48% respectively[18]. A more in-depth classification system that links prognosis and treatment pathways is more commonly seen and is discussed in section 1.2.6.

1.2.1 HCC epidemiology

1.2.1.1 *Incidence*

Diagnostic confirmation of HCC is not routine practice worldwide, however due to the large percentage of liver cancers being HCC, it can reliably be assumed that the incidence of primary liver cancer worldwide reflects the rate of HCC.

HCC is the 6th most common cancer worldwide, with 841,000 new cases in 2018. This accounts for 4.7% of all cancers diagnosed in 2018[1]. HCC is more common in men than women, ranking 5th and 9th respectively, with age-standardised rates 25% higher in men[1].

HCC distribution is greatly affected by geography, race and ethnicity. 76% of the 5-year prevalent cancer cases from 2008-2012 occurred in Asia[19]. China alone accounted for 50% of all new cases in 2012. HCC has very defined risk groups, with most cases having some form of underlying disease. The high proportion of cases in Asia is primarily accounted for by the rates of hepatitis B and C virus infections.

In contrast to rising rates in western regions, incidence rates in Asia and sub-Saharan Africa are on the decline. Falls in incidence rates coincide with the introduction of hepatitis B virus immunisation programs, while rates are increasing in more western regions alongside the rise in obesity and alcohol consumption. HCV related HCC's are also on the rise due to epidemic outbreaks caused by IV drug use[18] and periods of poorly sterilised medical equipment[18].

1.2.1.2 *Mortality*

HCC has one of the worst prognoses of all cancers. Despite ranking 6th for incidence rates in 2018, it ranked 4th in the number of deaths from any specific cancer[1]. The number of deaths in 2018 almost equals the number of cases,

782,000 and 841,000 respectively, giving an overall odds ratio of mortality to incidence of 0.93. Unlike many other cancers, the mortality rates are similar throughout the world. Odds ratios in more developed regions are 0.92, and 0.96 in less developed regions.

The poor prognosis seen in liver cancer is multifactorial. Liver cancer most often arises after development of liver fibrosis and cirrhosis, causing severe impairment of liver function. Due to the vital role of the liver in normal bodily function this leads to a high risk of mortality. The position of the liver in the body can make imaging techniques unsatisfactory for early diagnosis, this along with the paucity of early symptoms, leads to HCC being diagnosed at a late stage, often after the occurrence of metastasis. Due to this, complete resection is rarely possible and full curative liver transplantation is often unavailable due to lack of donor organs.

1.2.1.3 *Gender*

HCC is more common in men than women. Part of this is attributed to the higher rates of HBV and HCV infection seen in men, as well as a tendency to higher rates of chronic alcohol consumption. The primarily female hormone prolactin (PRL) has been shown to have a protective effect against HCC. Postmenopausal women not undergoing hormone replacement therapy (HRT) show an increased rate of liver cancer. Female $PRL^{-/-}$ mice had similar rates of tumour development to male $PRL^{-/-}$ mice when exposed to the carcinogen diethylnitrosamine (DEN)[20]. Male mice treated with a dopamine D2 receptor agonist (which increases serum prolactin to near female levels) showed 22% incidence of HCC compared to 100% incidence in wild type male mice when exposed to DEN. Hartwell et al. (2014)[20] were also able to show that male livers are a more proinflammatory environment than female livers, backed up the by the evidence that men with chronic HCV are more likely to progress to cirrhosis. Despite differences in risk, gender does not show an effect on mortality rates.

1.2.1.4 *Age*

As with most non-hereditary cancers, HCC is a disease of the ageing population. HCC is rare in the under 40s with a sharp rise in rates beyond that age group[19]. Another sharp rise is seen in the over 70s, with HCC cases in the over 70s almost equal to all cases in under 70s.

1.2.2 **High risk groups**

1.2.2.1 *Cirrhosis*

Liver cirrhosis is the primary risk factor of HCC with 80-90% of HCCs arising from underlying cirrhosis[21]. Cirrhosis is advanced un-reversible fibrosis and scaring of the liver. Minor fibrosis is reversible through reduction or removal of the factors leading to the fibrosis such as alcohol, high fat diets and viral hepatitis. This precursor damage to the liver partially accounts for the high mortality rates seen in HCC. The process of cirrhosis and subsequently HCC development is closely linked with the chronic inflammation triggered as a response to liver injury, and leading to activation of local immune cells (Kupffer cells, liver dendritic cells, T cells, natural killer T cells and natural killer cells) and subsequent recruitment of non-resident immune cells[22]. Chronic exposure to highly active immune cells results in production of a proinflammatory cytokine storm and ultimately resulting in hepatocyte cell death, commonly referred to as necroinflammation. In cases of sever liver damage, hepatocyte regeneration is impeded by cellular senescence or growth arrest[22]. Ductal reactions (DRs) act to preserve progenitor cell populations and forms regenerating nodules (RNs), which are a hallmark or liver cirrhosis[22]. The rapid rate of cellular proliferation in RNs can lead to accumulation of DNA damage, leading to dysplastic nodules and ultimately malignant transformation to early HCC. Figure 3 demonstrates the histopathological transformations taking place during HCC development in a background of cirrhosis.

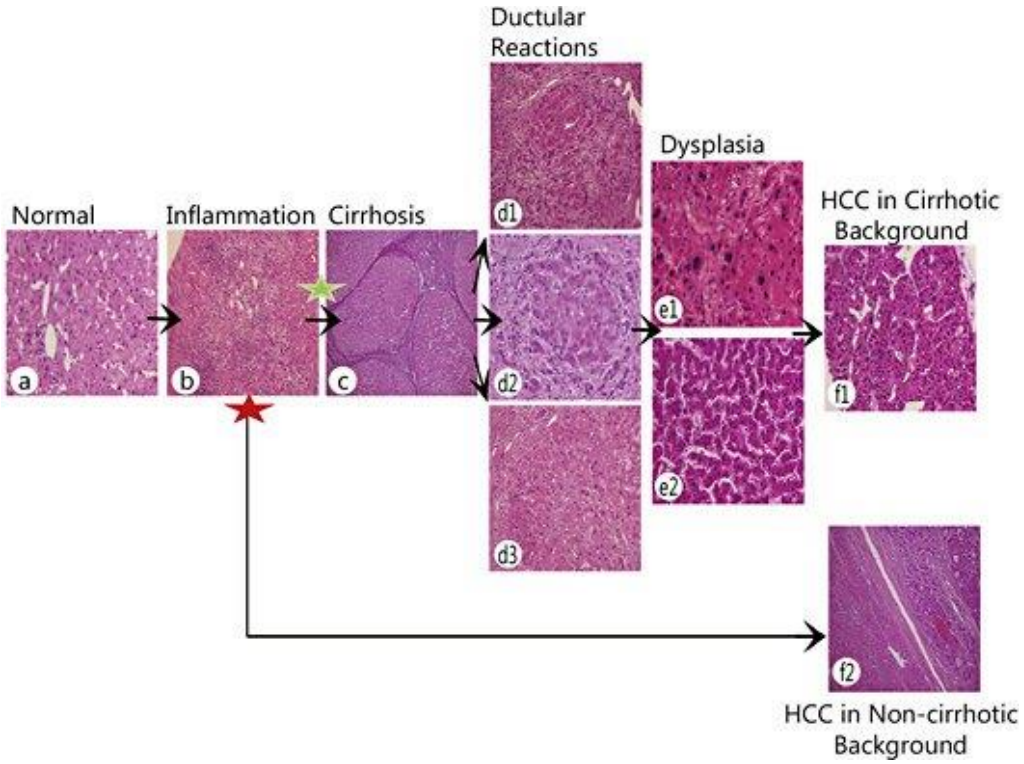


Figure 3 - Chronological representation of histopathological events during progression toward HCC. The majority of HCC lesions occur against a cirrhotic background (green asterisk), whereas in rare instances it can also occur against a noncirrhotic background (red asterisk). Typical histopathology sections during disease progression are shown: (a) normal liver (×200), (b) chronic hepatitis with marked portal/peri-portal chronic inflammation (×100), (c) cirrhosis with regenerating parenchymal nodules separated by active fibrous septa (×40), (d1) ductular reaction in cirrhosis (×100), (d2) ductular reaction with different stages of differentiation toward hepatocytes in a regenerating nodule (×200), (d3) large cell changes in regenerating nodule in cirrhosis (×100), (e1) large cell dysplasia (×200), (e2) small cell dysplasia (×200), (f1) HCC with thick trabeculae of malignant hepatocytes (×200) and (f2) HCC with trabeculae and pseudoacinar transformation in a noncirrhotic background (×40). Reproduced with permission from Ramakrishna et al. (2013)[22]

The liver processes blood from all digestive organs. When the liver is scarred, blood flow through the liver becomes blocked, leading to hypertension due to increased blood pressure in the portal vein. Portal hypertension causes blood to flow backwards leading to oesophageal and stomach varices[23]. Blood flow back to the spleen causes splenomegaly, which can lead to thrombocytopenia, anaemia and neutropenia. Hepatic venous pressure gradient has been linked to the development of HCC in decompensated alcoholic cirrhotic patients[24]. Partial Thromboplastin time test is used to gauge liver damage by assessing blood clotting ability. Scarring of the liver can be monitored by assessing liver stiffness via transient elastography scanning.

Cirrhosis can also cause hepatic encephalopathy, due to reduced ability to detoxify ammonia, therefore increasing brain levels of ammonia, resulting in confusion, loss of concentration, memory loss, coma and even death. Hepatic pulmonary syndrome occurs in around one third of cirrhosis cases. It is thought that advanced cirrhosis leads to secretion of hormones such as endothelin-1, a vasoconstrictor which causes smooth muscle to contract. In the absence of smooth muscle, such as in the lung's capillaries, it will cause vasodilation, by releasing nitric oxide from capillary cells, increasing blood flow. This causes too many red blood cells to flow through the alveoli at once, leading to insufficient oxygenation and a build-up of pressure.

Reversed blood flow due to portal hypertension causes abdominal ascites, which are a risk factor for infections. This also reduces the volume of blood that will be returned to the heart causing hypovolemia. Hypovolemia will cause arterial systemic vasodilation. As the kidneys are not receiving enough blood, they will activate the renin-angiotensin aldosterone system[25]. This will stimulate absorption of sodium and water from the kidneys therefore increasing water and sodium levels in the body to compensate for the lack of blood supply, causing oedema in the hands and feet.

Jaundice is the most visually identifiable sign of liver damage. Obstruction of bile flow from the liver into the gall bladder and small intestines due to scarring causes a rise in plasma bilirubin levels leading to yellowing of the skin.

As cirrhosis mainly develops in the background of other liver diseases, it is hard to pinpoint any molecular mechanisms of cirrhosis alone that lead to HCC formation. Telomere dysfunction often occurs in hepatocytes of a cirrhotic liver. Telomere dysfunction has been linked with an increase in early stage hepatic neoplasms; however, its effect is limited by cell cycle regulators, such as p53, p53 mutations in combination with telomere dysfunction lead to accelerated tumour onset[26].

1.2.2.1.1 Non-alcoholic fatty liver disease

Non-alcoholic fatty liver disease is an increasing risk factor of cirrhosis and HCC in western and rapidly developing countries[27]. NAFLD is primarily diet induced, with the rise in incidence linked to modern high fat/high fructose diets. Fat build up in hepatocytes, due NAFLD, is known as hepatic steatosis. NAFLD is closely linked with metabolic syndrome, particularly type 2 diabetes mellitus and obesity[28].

Steatosis occurs when the rate of uptake and synthesis of lipids is greater than oxidation and secretion of lipids. Hepatocytes secrete the enzyme hepatic lipase, which breaks down triglycerides into fatty acids to facilitate absorption[29]. In hepatocytes, fatty acids can be synthesised into triglycerides or vice versa. If there are high levels of fatty acids in the blood, they will be absorbed by the hepatocytes. This leads to an increase in triglyceride synthesis and reduction of fatty acid oxidation leading to steatosis. Fructose can also be metabolised into fatty acids by enzymes in the liver. Fructose consumption has generally increased with the use of high-fructose corn syrups in sweet products such as fizzy drinks.

Metabolic syndrome is a collective name for a group of metabolic diseases; central obesity, hypertension, dyslipidaemia, impaired glucose tolerance and insulin resistance, three of five of these diseases are enough to cause metabolic syndrome[30]. Metabolic syndrome tends to affect the pathways regulating fatty acid oxidation and conversion to triglycerides; therefore, increasing problems with fatty acid build up in hepatocytes.

Fatty acid build-up in hepatocytes is thought to increase oxidative stress. Lipid peroxidation products along with reactive oxygen species (ROS) lead to DNA damage as well as cytokine production including TNF-alpha, TGF-beta and FAS ligand. Excessive production of TNF-alpha and TGF-beta has been linked with the formation of multiple cancers[31,32]. Inflammation caused by cytokines leads to fibrosis development and is the major cause of Non-alcoholic Steatohepatitis (NASH) the most severe form of NAFLD[33].

1.2.2.1.2 Alcoholic liver disease

High alcohol consumption is a major risk factor for steatosis. This is due to the breakdown pathway of ethanol[34]. Ethanol is converted to acetaldehyde by alcohol dehydrogenase. Acetaldehyde is then converted to acetic acid by aldehyde dehydrogenase, and subsequently to acetyl-coA which can be converted to fatty acids. Breakdown of ethanol and acetyl aldehyde requires reduction of nicotinamide adenine dinucleotide (NAD). Excess reduced NAD ($NADH+H^+$) causes shunting of normal substrates away from metabolism and towards lipid synthesis. Due to these pathways alcoholics are a high-risk group for liver cirrhosis.

Alcohol also has a direct effect on HCC formation. Acetaldehyde is both toxic and mutagenic. Acetaldehyde causes the DNA adducts N^2-ethyl-2'-deoxyguanosine (N^2-ethyl-dG)[35] and 1,N(2)-propano-2'-deoxyguanosine (PdG)[36], leading to errors in DNA replication. During excessive alcohol intake, a second enzyme CYP2E1 breaks ethanol down to acetate then to acetaldehyde. CYP2E1 induction causes increased ROS generation, this results in lipid peroxidation producing other reactive aldehydes which can form DNA crosslinks and DNA-protein links[37]. Alcohol metabolites have been shown to induce collagen synthesis in hepatic stellate cells contributing towards fibrosis[38].

1.2.2.1.3 Wilsons disease

Wilson's disease is an inherited disease caused by a mutation in the *ATP7B* gene, responsible in regulating copper transport[39]. Copper is normally absorbed from the stomach and comes through the liver to be secreted in bile. The mutation leads to copper accumulation in the liver and the brain causing liver damage and neuropsychiatric symptoms. The liver damage contributes to inflammation and eventually cirrhosis.

1.2.2.1.4 Hereditary hemochromatosis

Hereditary hemochromatosis (HH) is caused by a mutation in the *HFE* gene which leads to an accelerated rate of iron absorption from the intestines, with

excessive iron deposited in tissues including the liver. HH causes cirrhosis along with hypopituitarism, cardiomyopathy, diabetes, arthritis and hyper pigmentation. HH is associated with an increased hazard ratio for HCC; although the ratio varies substantially between studies, almost all studies show an increased risk compared to healthy controls[40].

1.2.2.1.5 Autoimmune hepatitis

Autoimmune hepatitis (AIH) is an autoimmune disorder that can occur at any point and is more common in women than men. Inflammation and destruction of liver cells by the immune system leads to quick progression to cirrhosis. Although there seems to be no direct link to HCC, the development of cirrhosis causes increased risk of HCC. Risk of HCC among AIH patients is 1.9% per year, comparable to the average risk of HCC in all cirrhotic patients[41].

1.2.2.1.6 Hepatitis B

Chronic hepatitis B virus (HBV) infection is the primary individual risk factor for hepatocellular carcinoma[42], accounting for approximately 50% cases worldwide. HBV is estimated to have infected approximately two billion people worldwide and 90% of adults infected with HBV will suffer from acute infection, where they may or may not experience symptoms[43]. This 90% will recover and develop a lasting immune response preventing further infection. The remaining 10% will go on to develop chronic infection (virus remaining in the blood for >6 months). The outlook is not as good for children aged less than five years old, with 90% of those infected progressing to chronic HBV infection[43]. Chronic HBV infection is linked to a 5 to 15 fold increase in risk of HCC compared with the normal population[21].

HBV is transmitted through blood and other bodily fluids, including vertical transmission from mother to child. As both acute and chronic Hepatitis B can be asymptomatic, the virus is often unknowingly transmitted. HBV infection is diagnosed using blood tests for the presence of Hepatitis B surface antigen (HBsAg), this can be detected in both acute and chronic infections[44]. Hepatitis B e antigen (HBeAg) test monitors the level of HBeAg in the blood as a measure

of active infection; this signifies a high viral load and can be used to monitor treatment[44]. Chronic infection is diagnosed when either HBsAg or HBeAg is still detectable after six months.

The HBV genome (Figure 4) encodes just four open reading frames (ORFs), C, P, S and X. ORF C encodes the Core protein (HBcAg or HBeAg) which surrounds the HBV viral DNA and the HBV DNA polymerase encoded in the P ORF[45]. The S ORF encodes three surface proteins, preS1, preS2 and S (HBsAg), which form the viral envelope along with host cell membrane. This surface antigen has been shown to have multiple functions in supressing the immune response[46]. The full function of the hepatitis B x antigen (HBxAg), encoded for by the *X* gene, is not yet fully understood but has been shown to have critical roles in transactivation of cellular and viral genes vital for viral replication[47].

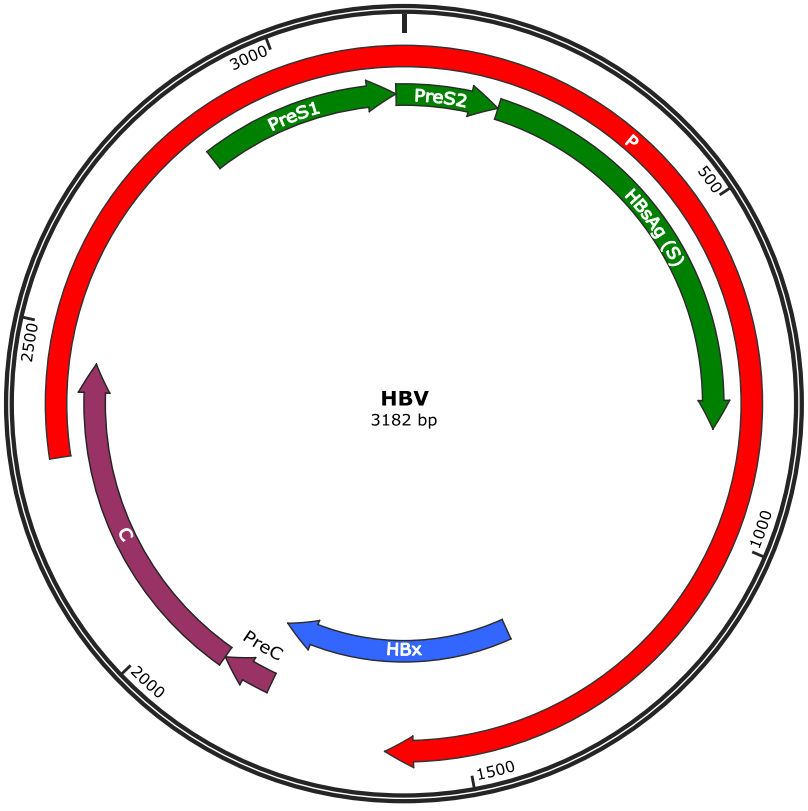


Figure 4 - HBV genome based on the reference sequence NC_003977.2

Hepatitis B can contribute to HCC either indirectly or directly. Indirectly, HBV production leads to death of the hepatocytes and increase of cell turnover. High rates of cell replications have a strong link with all cancers, as DNA replication can be error prone and allows the incorporation of mutations into

the cell genome, a process even more likely when considering the high levels of toxin exposure seen by hepatocytes (section 1.2.2.2.1).

1.2.2.1.7 Hepatitis C

Hepatitis C is the second most common individual risk factor of HCC after hepatitis B (as cirrhosis is caused by multiple risk factors). A study of 12000 Taiwanese men showed a 20 fold increase in risk of developing HCC in hepatitis C virus (HCV) infected subjects[48]. Approximately 170 million people worldwide are chronically infected with HCV[49]. Patients infected with HCV have a 2% annual risk of HCC and a 7-14% five year risk of HCC[50].

HCV infection, like hepatitis B virus (HBV) infection, remains asymptomatic and can be spontaneously cleared by the immune system. HCV, however, has a much higher chance of becoming a chronic infection than HBV with 55-85% of patients experiencing persistent infection[51]. There is no vaccine available for HCV, so treatment of infected individuals is the only option. HCV was classically treated with interferon and ribavirin with the additions of boceprevir and telaprevir for the classically difficult to treat genotype 1 infections. The addition of boceprevir and telaprevir brought treatment success rates of genotype 1 in line with other genotypes at 70-90%[52]. Recent trials of directly acting antivirals such as Daclatasvir and Sofosbuvir have shown treatment success rates of around 90% including patients unresponsive to previous treatments and have now become the standard treatment pathways in most countries[52]. Whilst this has been shown to reduce overall HCC incidence, the rate of HCC in patients cleared of HCV is still higher than the general population[53]. HCV linked HCC almost always occurs in a background of cirrhosis with chronic inflammation caused by increased cell turnover. HCV does not integrate into the genome as in HBV. Some HCV proteins have been linked with carcinogenic effects. The HCV core protein has been shown to interact and supress tumour suppressor proteins p53, p73 and RB, all of which are crucial to preventing cancer formation[54]. It has also been shown to affect the RAF1/MAPK pathway leading to increased TGF-beta expression[55], and its interaction with SMAD3 switches TGF-beta's tumour suppressive effects to tumour inducing. The HCV E2 protein

inhibits natural killer (NK) cells helping the virus to evade the immune system as well as suppressing the anti-cancer abilities of the immune system[56]. The NS3 proteins can supress p53 and p21 activity as well as inducing TNF-alpha leading to inflammation[57]. NS5a is able to block apoptosis via interaction with the binding domains of the pro-apoptotic protein BAX[58].

1.2.2.2 *Non-Cirrhotic HCC*

1.2.2.2.1 Hepatitis B

As well as causing HCC through cirrhosis, hepatitis B can also cause HCC through non-cirrhotic routes. The HBx gene encodes an 154 amino acid (AA) protein, HBxAg, which has been shown to have multiple effects with direct links to HCC formation. The HBxAg has been demonstrated to stimulate cancer linked cell proliferation pathways including the Ras-Raf-MAP kinase pathway. It accomplishes this through protein-protein interactions with cellular transcription. The HBxAg has been linked with increased production of reactive oxygen species (ROS) leading to DNA damage and mutation accumulation[59] as well as changes in DNA methylation through the transcriptional and translational regulation of DNA methyltransferases[59]. Incorporation of HBV DNA into human DNA often leads of point mutations and C-terminal truncation of the HBx gene. These mutations are linked to increased cellular proliferation, enhanced hepatoma cell invasion and metastasis[60].

Ayub et al. (2013)[61] published an extensive review of proteins misregulated in HBV infection, finding 31 up-regulated and 18 down-regulated. The majority of these proteins are primarily involved in cell proliferation pathways, such as downregulation of the crucial cell cycle controller p53. Other pathways affected include immunosuppression and protein synthesis.

1.2.2.2.2 Aflatoxin

Aflatoxin B1 (AFB1) is considered to be one of the most potent carcinogens. It is produced by the *Aspergillus spp*. fungus primarily found in Asia and sub-Saharan Africa. The primary method of aflatoxin exposure is through contaminated agricultural products including rice, peanuts, cereals and grains.

It is estimated that 4.5 billion people are regularly exposed to aflatoxin through food, especially in low income countries[60]. AFB1 is metabolised by the liver into AFB1-8, 9-epoxide[7], which can intercalate DNA and alkylate bases. AFB1 exposure is linked with a mutation in the *p53* gene in patients with HCC[62].

1.2.2.3 *Diabetes Mellitus*

An increased risk of HCC in patients with diabetes mellitus has been shown in multiple studies[63]. Pre-existing diabetes is associated with adverse outcomes in hepatocellular carcinoma, from occurrence, progression, and to mortality. While the mechanism underlying the association between diabetes and prognosis remains unclear[63].

1.2.3 **Molecular pathogenesis**

Most cancers are heterogenous, however the extent of this in HCC is much greater than most. For example, almost all breast cancers can be divided into one of five classical subtypes (luminal A, luminal B, HER-2, basal and normal breast like) based on the underlying molecular differences[64,65]. These subtypes also correlate well with prognosis and have allowed the development of specific therapies to target pathways affected in these subtypes such as Trastuzumab (Herceptin) which has been developed to treat HER-2 positive tumours[66].

HCC appears to develop as an accumulation of various mutations from several pathways as appose to a small set of common mutations. Whilst this is the case, many mutations are seen in high percentages of HCC patients however with varying combinations.

Telomerase promoter mutations have been shown to be the most common mutations in HCC, being present in 30-60% of HCCs[67]. Extensive telomere shortening leads to activation of DNA damage pathways which trigger apoptosis or cellular senescence and is linked with the high cellular turnover caused by viral and non-viral damage. Telomere shortening results in chromosomal instability and may drive cancer initiation. Activating mutations or HBV viral DNA insertions of the TERT promoter are key for cellular

immortalization, allowing cells to survive and proliferate indefinitely[67]. TERT mutations are seen in early progression to HCC, being found in 6% of low-grade and 19% of high-grade dysplastic nodules[68].

TP53 mutations are abundant in almost all common malignancies. P53 is a key cell cycle regulator which cancer cells must overcome to proliferate. P53 is mutated in more than 50% of aflatoxin B1 HCCs, 45% of HBV-related HCCs and 13% of HCV related HCCs[69]. Single base substitutions of P53 occur in approximately 25% of HCCs, primarily affecting its DNA binding domain[70]. These mutations effect binding to the induction region of P53's primary regulator, MDM2 [71],as well as downstream cell cycle regulators and apoptosis inducers. Arginine to serine mutations at codon 249 are the most common mutations seen in aflatoxin B1 related HCC[72].

The canonical Wnt signalling pathway is affected in 20-90% of HCCs[73]. β-catenin mutations which lead to accumulation are present in 20-40% of HCC and mutations to AXIN1 and AXIN2, part of the β-catenin degradation complex, are present in 3-16% of HCCs[74]. β-catenin mutations are associated with increased micro and macro-vascular invasion as well as upregulation of glutamine synthase[75]. Direct mutations to β-catenin are seen more commonly in HCV related (33.3%) and alcohol related (41.8%) than in HBV related (11.4%) HCCs[74]. However, HBV is associated with Wnt pathway disruption through inhibition of the β-catenin degradation by the HBx protein[76].

Copy number and chromosomal alterations are common in HCC as with other solid tumours. A study of 125 HCCs identified focal amplifications in 32% of HCC and focal deletions in 40%, with FGF19 and CCND1 commonly amplified and CDKN2A/B, AXIN1 and IRF2 commonly deleted[77]. Similar mutations were found in another study of 286 HCC patients, with high level copy number changes in CCND1, CDKN2A/B and MET[78]. Less frequently affected pathways include chromatin remodelling (through ARID1A/ARID2 mutations), PI3K/Ras and the oxidative/ER stress pathways. PIK3CA mutations highlight the variations seen by region, with mutations occurring in 35.6% of HCC cases in China[79], 28% of cases in Italy[80], but 0% of cases in Japan[81].

Many of the discussed molecular alterations in HCC are not easily druggable, and has led to a lack of agents available to treat HCC[82]. Sorafenib and regorafenib are the only non-immunotherapy agents approved to treat HCC. Both sorafenib[83] and regorafenib[84] are multi-kinase inhibitors which target the tumour cell's serine-threonine kinases RAF/MEK/ERK, which are involved in tumour cell progression, as well as receptor tyrosine kinase angiogenesis promotors VEGFR1/2 and platelet derived growth factor receptor (PDGFR). The multi-target action of these drugs hints towards their success in a highly molecularly variable disease such as HCC. Attempts to enrich trials of new drugs based on molecular subtyping have so far failed to show benefits[82], for example the phase III METIV-HCC trial of tivantinib, a cMET inhibitor, targeted patients with elevated cMET expression determined by immunohistochemistry that had progressed on sorafenib, showed no significant advantage in overall survival[85].

1.2.4 **Screening**

Due to the uncertainty over current markers, there is no nation screening strategy for HCC in the UK. Current AALSD recommendations are for ultrasound (US) screening in only the highest risk HCV and HBV patients in the USA[86]. Alpha-fetoprotein (AFP) has been approved for use in China and Japan, with a combination of AFP and des-gamma-carboxy prothrombin (DCP) measurement covered by the Japanese nation health insurance[87]. The largest randomised control trial (RCT) for screening of HCC analysed 18,816 patients from China with HBV infection or history of chronic hepatitis[88]. Whilst this study has attracted criticism for its statistical analysis, it showed a clear reduction in stage at diagnosis and mortality rates. During this study patients received an AFP test and ultrasonography every six months. HCC screening has also been shown to reduce stage at presentation, especially in studies which correct for lead time bias[89,90]. Reduction in stage at presentation has a clear benefit in overall survival of HCC, however this is not the only outcome required to implement screening plans. Health economic studies must also be conducted to assess the feasibility of introducing such plans. A systematic review of HCC screening articles concluded that HCC screening is possible at a reasonable cost per

quality adjusted life year gained (QALY), however they stress the need for an appropriate RCT to confirm this[91].

The following methods for HCC screening are summarised in Table 1. It is important to note that these figures compare the effectiveness at detecting all stages of HCC and are often significantly lower for early stages tumours.

Table 1 – Comparison of Area under the receiver operating characteristic curve (AUC) of screening tests for HCC for all stages of tumour.

Marker	Reference	AUC
AFP	[92]	0.83
DCP	[92]	0.79
AFP-L3	[92]	0.71
OPN	[93]	0.78
GPC3	[94]	0.76
GP73	[92]	0.91
GALAD	[95]	0.95
Ultrasound	[96]	0.96

1.2.4.1 *Ultrasound*

Ultrasound is the primary screening tool for HCC worldwide. When used as a screening tool, US is typically repeated every six months due to the doubling time of HCC[97]. Ultrasound suffers from a poor ability to distinguish early stage HCC. Cirrhosis and fat can both interfere with imaging causing variations in the image seen, leading to difficulties in detecting lesions less than 3cm. Due to variation in image output, operator skill also contributes to a large variation in performance. A meta-analysis by Singal et al. (2009)[98] showed sensitivities ranging from 23%-98% , in the majority of these studies, combination with AFP showed no additional increase in sensitivity.

1.2.4.2 *Alpha-fetoprotein*

Alpha-fetoprotein (AFP) is a 70kda protein expressed at high levels in the human foetus, which drop dramatically to undetectable adult levels within 12 months of birth. AFP screening is used as a predictor of foetal chromosomal abnormalities in the third-trimester. AFP was first shown in the serum of HCC patients in 1964[99]. Since then, studies have shown conflicting data to support

its use as a screening tool. It is still unclear whether AFP in combination with US is an effective enough strategy to reduce mortality. AFP occurs in the serum of many patents with non-cancerous chronic liver disease (NCCLD), therefore struggles for specificity. Cut-offs between 20ng/ml-400ng/ml have been used in various studies causing difficulties in assessing inter-study performance[100]. Sensitivity for all stages of disease commonly ranges from 41-65% with specificity ranging from 80-94%, [100]. However, the sensitivity for early stage detection is typically below 50%[101,102].

1.2.4.3 *Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein*

AFP can be fractioned into three glycoforms; L1, L2 and L3. Lens culinaris agglutinin reactive fraction of alpha-fetoprotein (AFP-L3) is a fucosylated form of AFP, is thought to be produced only by cancerous cells, whereas AFP-L1 is often produced in liver cirrhosis[103]. Its use as a test for HCC assesses the ratio of AFP-L3 to total AFP. A meta-analysis of AFP-L3 trials shows an increased specificity of around 93% and a sensitivity of 48.3% for all stages of HCC[104]. AFP-L3 still requires further trials to assess its validity as a marker of HCC.

1.2.4.4 *Des-gamma-carboxy prothrombin*

Des-gamma-carboxy prothrombin is the modified form of prothrombin induced by vitamin k2 absence/agonist II (also known as PIVKA-II). It completely lacks the normal coagulation activity of prothrombin and has been linked to malignant proliferation of HCC[105]. As with many developing markers for HCC, performance has varied greatly between studies[106]. Part of this is due to kits produced by different manufacturers and various cut-off levels used. Pooled results from a meta-analysis of 12 studies found the sensitivity and specificity to be 71% and 84% respectively[106]. This is similar in performance to AFP. Whether or not DCP is complimentary to AFP measurement is still debated, however Japanese guidelines recommend screening with AFP and DCP.

1.2.4.5 *Golgi protein 73*

Golgi protein 73 (GP73) has shown similar accuracy to AFP in the diagnosis of HCC with sensitivities ranging from 7% to 91% and specificities ranging from 52% to 97% with pooled sensitivity and specificity of 76% and 86% respectively[107]. Studies are ongoing to look at the potential of combining it with AFP to produce a more powerful test.

1.2.4.6 *Osteopontin*

Osteopontin (OPN) is a multifunctional calcium binding-phosphorylated-acidic glycoprotein of the extracellular matrix (ECM) that is overexpressed in several malignancies as well as increased rates linked to inflammatory processes[108]. Shang et al. (2011) found OPN to have high sensitivity for HCC (91%) but low specificity in cirrhotic and chronic HBV patients (61%)[108]. This is likely related to the high inflammation seen in the livers of patients with cirrhosis and chronic HBV infected patients. As OPN is overexpressed in other malignancies, OPN positive patients must be followed up with imaging to detect the location of the malignancy.

1.2.4.7 *Glypican-3*

Glypican-3 (GPC3), similar to AFP, is expressed in the liver during embryonic development however is not seen in the liver of healthy adults[108]. GPC3 has been demonstrated to aid in tumourigenesis via activation of the Wnt signalling pathway[108]. A meta-analysis of serum GPC3 levels in HCC patients vs controls, found a pooled sensitivity of 55.2% and specificity of 84.25% and combination with AFP significantly increased sensitivity with no significant effect on specificity[94].

1.2.4.8 *Micro-RNAs*

Micro-RNA (miRNA) analysis is an emerging field in diagnostics. A systematic review and meta-analysis identified 17 studies from 2010-2014 investigating serum miRNA levels in HCC patients[109]. The most commonly reported miRNAs with altered expression in HCC were miR-21, miR-122 and miR-233, with AUCs of 0.9293, 0.8128 and 0.8597 respectively, for distinguishing HCC from healthy

controls. These results look promising; however, most markers of HCC fall short in distinguishing HCC from NCCLD, an issue considering 90% of HCCs arise in patients with liver disease. The best performing miRNA, miR-21, is not a specific marker of HCC with altered levels detectable in multiple cancers[110]. Whilst studies have shown potential for miRNA use as biomarkers of cancer, there are several pitfalls including sampling inconsistencies, sample instability, non-standard measurement techniques and influence from a variety of factors including diet, smoking and environmental effects[111,112].

1.2.4.9 *GALAD model*

The GALAD model is a prediction score combining AFP, AFL-L3, DCP, age and gender. This score has shown improved sensitivities over AFP alone with minor reductions in specificities, leading to improvements in AUC. This has been demonstrated in sample sets from multiple countries and shows improved performance over AFP alone, however performance varies greatly depending on the region[113]. AFP, DCP and AFP-L3 measurements alongside ultrasound are all recommended in Japan every 3-4 months for HCC screening in the highest risk groups[114].

1.2.5 **Diagnosis**

Although US is often used in screening due to its low cost, computed tomography (CT) and magnetic resonance imaging (MRI) have replaced it as diagnostic imaging methods due to greater accuracies. HCC is difficult to detect via imaging methods as scarring caused by cirrhosis makes it difficult to detect underlying abnormalities. Cirrhosis also creates lesions that appear similar to a tumour leading to false positives[115].

CT and MRI are performed with the addition of a contrast agent. HCC tumours tend to derive their blood flow from the hepatic artery, whereas normal hepatocytes receive around 75-80% of their blood flow through the portal vein. This leads to contrast agent entering the tumour much quicker than the surrounding tissue. Contrast agents enters the tumour within 2-40 seconds of administration, and entering surrounding tissue between 50-90 seconds post

administration[116]. Small HCCs are difficult to detect through CT and MRI as most of their blood supply is derived from the portal vein. CT arteriography can increase detection rates, however due to its more invasive nature and added cost it is rarely used[117]. MRI imaging using superparamagnetic iron oxide (SPIO) particles is highly sensitive in detection of HCCs >1cm, however still has poor sensitivity in HCCs <1cm. Kupffer cells in the lobules take-up SPIO reducing

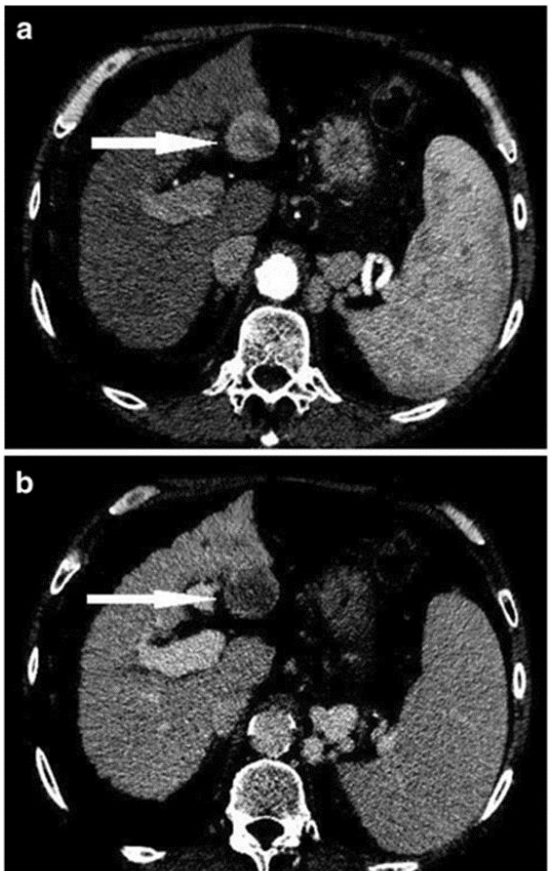


Figure 5 - 54 year old male with hepatitis C cirrhosis. CT shows an arterial enhancing nodule a with washout of contrast in the delayed phase b consistent with hepatocellular carcinoma. Willatt et al. (2017) [119]

signal intensity. This does not occur in HCC, but does still occur in liver disease and regenerating liver[118]. Key features of HCC in CT and MRI imaging are late arterial phase enhancement and contrast agent washout relative to the liver parenchyma during the venous or delayed phases (Figure 5)[119].

Biopsy remains the reference standard diagnostic test. Fine needle aspiration (FNA) guided by CT or MRI shows a specificity and sensitivity of 96% and 95% respectively[120]. Biopsy may not be necessary in cases where HCC diagnosis from imaging is certain. Small well-differentiated HCCs may be difficult to distinguish

from non-cancerous liver masses, where histological appearances may only slightly differ from normal hepatocytes. The most recognisable histological feature of pre-malignancy is dysplasia[117]. Biopsy is important in reducing false positive results from imaging before subjecting patients to liver resection or transplant.

1.2.6 **Staging**

The TNM staging system used widely in oncology is ineffective in HCC. The TNM staging system describes tumour size, lymph node metastasis and distant tissue metastasis[121]. This is not adequate for HCC, as the underlying state of liver function is often as large a concern as the malignancy. The Barcelona-Clinic liver cancer (BCLC) staging system is widely accepted as the best classification system of HCC and is endorsed by the European Association for the Study of the Liver (EASL), the European Organisation for Research and Treatment of Cancer (EORTC) and the American Association for the Study of Liver Diseases (AASLD)[122]. This system accounts for tumour number, tumour size and liver function, as well as an accompanied recommended treatment strategy for each classification (Figure 6).

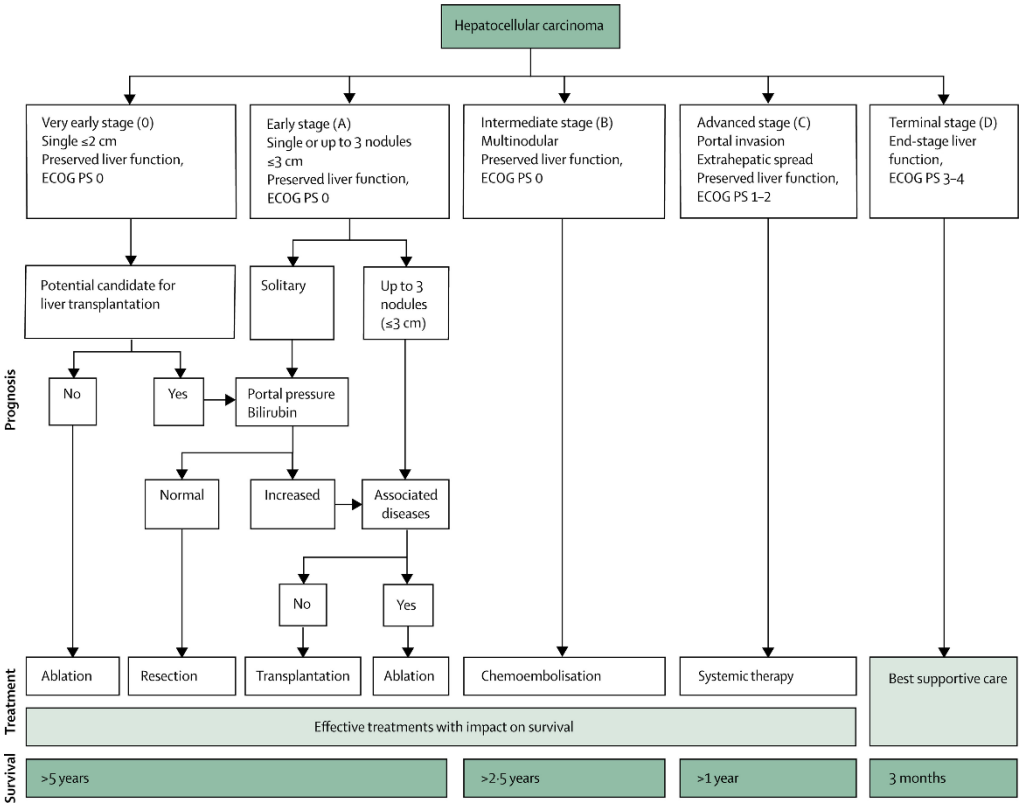


Figure 6 - BCLC staging system. Modified from Forner, Reig & Bruix (2018)[119]

1.2.7 **Treatments**

Recommended treatment options are dependent on BCLC stage at diagnosis. Curative resection is the preferred treatment choice however this is dependent on the underlying liver function. Many patients will have severe liver damage to the point where surgical resection will severely compromise remaining liver function. Resection is usually recommended in BCLC stage 0 patients[123]. Patients with low tumour burden but poor liver function will be of greater benefit from liver transplantation; however, donor availability is a major stumbling block. These patients usually fall under BCLC stage A. The Milan criteria recommends patients with one tumour <5cm or up to three tumours of <3cm with no macrovascular involvement or extrahepatic disease as eligible for transplant[124]. In the USA, the united network for organ sharing (UNOS) rank patients on the waiting list for transplant by their model for end stage liver disease (MELD) score. The MELD score looks at likelihood of death within the next 90 days due to liver function and assigns a number between 1 and 40, a

high score meaning higher chance of death. HCC patients that meet the Milan criteria tend to have a lower score than patients on the transplant list for NCCLD alone due to better overall liver function[125]. However, HCC patients' risk of death is often higher due to tumour progression rather than liver failure. To combat this, HCC patients meeting the Milan criteria are given a baseline score of 22[126]. BCLC stage A patients unsuited for resection/ liver transplantation are offered local ablation treatments, primarily radio frequency ablation or percutaneous ethanol injection.

BCLC stage B patients are often past the point of curative treatment. Transcatheter arterial chemoembolization (TACE) is the recommended treatment. These patients have extensive multifocal tumours without vascular invasion but preserved liver function. Two year survival rates in these patients are poor at 41%[127].

BCLC stage C are considered advanced HCC patients and are typically recommended for systemic treatment. The primary drug choice is sorafenib, however recent trials have shown efficacy of regorafenib in these patients[128]. Patients progressing on these treatments are usually offered into trials to assess new anti-tumour drugs including immunotherapy agents, such as nivolumab, an inhibitor of PD-1, which has recently achieved FDA approval as a second line therapy after sorafenib[128].

BCBL stage D are considered terminal. These patients receive palliative care for symptomatic relief and to ease suffering.

1.3 Immune system

The primary role of the immune systems is to fight off and protect our body from infectious agents. It accomplishes this using two systems, the innate and the adaptive system. The innate immune system produces a rapid nonspecific response that is localised to the site of infection. The adaptive immune system is a more advanced response, able to provide lasting immunity to infections. The adaptive immune system primarily comprises of T and B cells. T and B cells are stimulated by the recognition of an antigen usually thought of as non-self.

This triggers differentiation of T and B cells and initiates an immune response. T Cells mature into CD8+ cytotoxic cells or CD4+ helper cells. CD8+ cells act to directly kill cells presenting an antigen through major histocompatibility complex class 1 (MHC1) which is recognised by the T cell receptor. CD4+ cells produce cytokines which attract cells to deal with specific pathogen types, such as eosinophils in parasite infections. B cells contain an immobilised antibody molecule as a receptor. These antibody-receptors recognise antigens in their native form, rather than the processed form in MHC1 and 2 molecules. Once a B cell recognises an antigen it will differentiate into short lived plasma cells which secrete antibodies that target that antigen, or develop into memory B cells[129]. Memory B cells remain in the germinal centre undergoing affinity maturation. Some plasma cells become long-lived plasma cells (LLPCs) and remain in the bone marrow. A combination of LLPCs and memory B cells creates immune memory. These antibodies facilitate phagocyte recognition and trigger the complement cascade. Antibodies can block active sites of antigenic molecules and mark them for phagocytosis. Antibodies are also able to block the function of antigenic molecules by aggregating the antigen in immune complexes therefor preventing their function due to blocking of active sites and reducing ability to penetrate susceptible cells[130].

T cell and B cell responses are predominantly dependent on binding to antigens. What makes a good antigen differs depending on the response. T cell responses rely on denatured, unfolded peptide presentation by MHC molecules and is separated into two major processes. MHC1 molecules present peptide sequences to CD8+ T cells, whilst MHC2 molecules present peptides to CD4+ T cells. MHC1 molecules are produced in most cells and present intracellular antigens, whereas MHC2 are produced in specialised cells and present exogenous antigens, however some degree of cross-presentation exists between the two pathways[131]. MHC1 molecules present peptides which are generated by the proteolysis of proteins tagged via ubiquitylation by the 26S proteasome[132]. Degraded peptides are translocated to the endoplasmic reticulum where a MHC1 molecule is assembled from a polymorphic heavy

chain and an invariant light chain (β2-microglobulin) with the peptide inserted into the MCH1 binding groove[133]. Upon peptide binding, chaperone proteins are released from the MHC1 molecule allowing transportation to the cell membrane. What creates a good antigenic peptide for MHC1 molecules is less determinant on the peptide and more so the ability of the individual's HLA allelic variance which causes unique differences in MHC1 peptide grooves[134].

MHC2 is primarily expressed by specialised antigen presenting cells such as dendritic cells, macrophages and B cells. MHC2 molecules contain 2 heavy chains rather than a heavy and light chain as in MHC1. MHC2 molecules are assembled with a class II-associated Li peptide in the binding grove, which is later exchanged for peptides derived from proteins degraded during endosomal take-up[135]. After association with a peptide the MHC2 molecule transports to the cell membrane to present the peptide to CD4+ T cells.

B cells recognise whole antigens via their B cell receptors (BCRs). BCRs are formed in a similar process to mature antibodies and are almost identical to IgM and IgD with the addition of a transmembrane domain and CD79a/b cytoplasmic domains[131]. V(D)J rearrangements generate a vast diversity of immunoglobulin variable heavy chains (IgH) which then pair with VJ variable immunoglobulin light chains (IgL) to form the pre-BCR. These pre-BCRs are probed for autoreactivity before being displayed as a BCR on the cell membranes of immature B cells, with each cell expressing only one BCR. Immature B cells migrate from the bone marrow to secondary lymphoid tissues, primarily the spleen and lymph nodes, for maturation (germinal centre reaction)[136]. B cells can also be attracted to tertiary lymphoid structures formed at sites of chronic inflation such as in infections, autoimmune diseases and tumours[137]. Lymphatic fluid supplied to the lymph nodes contains a sample of soluble and cell bound antigens found in interstitial fluid and provides an opportunity for the BCR to bind its cognate antigen. BCRs bind either small antigens or peptides present in the lymphatic fluid or larger antigens which are presented in either a processed or native form by macrophages and dendritic cells[138]. Bound BCRs internalise and processes the antigen for membrane

presentation before movement towards the T cell zone of the lymph nodes where they can receive signals from CD4+ T cells via MHC2 binding of presented peptides[139]. These B cells then undergo class switching and maturation, triggered by binding of the T cell CD40 ligand to the B cell CD40 receptor. to form extrafollicular plasma cells producing low affinity antibodies or undergo affinity maturation to produce high affinity anitbodies[138]. T cell maturation is also triggered by binding of B cell ligands such as B7 to T cell receptors such as CD28[139]. It has been shown that T cells and B cells working in co-operation do not necessarily recognise the same epitope of the same antigen[140].

There have been many attempts to determine what makes a good BCR/antibody epitope, however these have been moderately successful at best[141]. BCR epitopes can be discontinuous, requiring proper folding of the antigen, or continuous, which are linear epitopes of continuous amino acids. It is thought that linear epitopes correspond to unfolded regions of proteins rather than their native forms[142]. Studies to identify natural binding of discontinuous epitopes are often flawed by the observation that binding of proteins in solid phase immunoassays to plastic cause a denaturing effect of the protein and is evident in many monoclonal antibodies which are suitable for use in denatured techniques such as western blot, however, do not function in native techniques such as immunohistochemistry. Much of the knowledge of what makes a good antigen/epitope come from studies based on producing research monoclonal antibodies and vaccine development. Primarily antigens must display a degree of 'foreignness' as negative selection takes place in the development of B cells to remove those which recognise self-proteins to limit autoimmunity[143]. Immunogenicity is correlated with the biological complexity of the substance, with proteins being generally more immunogenic than nucleic acids and polysaccharides[144]. Areas of structural stability are key for B cells responses as antibodies primarily bind in a conformational manner, and therefore their binding can be prevented if confirmation of the target molecule changes. By a similar manner, epitopes are often found in the terminal regions of proteins as these are more likely to be exposed to the BCR/antibody[145,146],

whereas central molecules are more likely to produce peptides which can be presented through MHC molecules. Larger molecules and higher exposure increase the chance of recognition by a receptor/antibody[147].

Whilst these are common features of non-native antigens, the presence of auto-reactive antibodies in caner shows that these are not the only factors present in an epitope. Changes to proteins during the development of cancer may however cause self-proteins to acquire more of these features. To that extent, sections 1.5.1-1.5.4 represent common factors trigger autoantibody production in cancer.

1.4 Autoimmunity and Autoantibodies

It has been shown that T and B cells can also recognise self-antigens[148]. This is apparent in autoimmune disorders, where immune cells recognise self-antigens and mount an immune response against host cells. There are over 80 different autoimmune disorders that we are currently aware of, often characterised by the presence of autoantibodies with certain specificities that can cause long lasting disease and debilitating conditions. These vary in severity from rashes seen in psoriasis to complete organ damage seen type 1 diabetes and in severe conditions leading to death. Systemic lupus erythematosus occurs with the production of several AAbs including anti-nuclear, anti-dsDNA, anti-smith, anti-Ro, anti-LA and anti-U1 RNP antibodies, and affects connective tissue throughout the body, causing multiple organ damage. Like many autoimmune disorders it can flare up and then go into periods of remission. Autoimmune disorders can be caused by a wide range of factors including genetic susceptibility, environmental factors, reactions to drugs and hormonal changes[149]. Many autoimmune disorders appear in children going through puberty and women of child bearing age.

Minor autoimmunity however can have beneficial effects. Autoimmune response against epithelial cells in response to stress or damage is thought to play a role in tissue homeostasis[150]. This includes the damage and stress caused to cells by cancer.

1.5 Autoantibodies in cancer

Autoantibody (AAb) production has also been described in cancer patients and is thought to be triggered by the proinflammatory nature of tumour establishment and growth[151]. These AAbs are primarily targeted at tumour associated antigens (TAAs)[151]. Cancer can be recognised by the immune system in a number of ways[152]. Stress related ligands MICA and MICB are often produced by cancer cells. These ligands can be recognised by the pattern recognition receptors of NK cells, phagocytes and cytotoxic t-cells. NK cells can also be activated by the common loss of MHC1 molecules from cancerous cells[153]. The adaptive immune system is often triggered by dendritic cells which have phagocytised tumour debris and presents them using its MHC1 molecules[154]. This stimulates CD8+ cell response against the tumour and CD4+ mediated B cell antibody production. There are three primary forms of TAA that are recognised: mutated proteins, aberrantly expressed proteins and post-translationally modified proteins. It is thought that this process helps to prevent development of many tumours during an individual's lifetime, and those tumours which obtain the ability to evade this process are those which go on to develop and become clinically relevant, however the exact role of AAbs, and the immune system as a whole, in tumour control and progression is still debated[151]. On the contrary, evidence exists which shows a pro-tumourigenic effect of B cells and associated AABs. Circulating immune complexes of AAbs and antigens can induce myeloid derived suppressor cells, through FCγ receptor interaction, which supress anti-tumour CD4+ and CD8+ T cell responses[155].

1.5.1 Mutated TAAs

Anti-p53 AAbs are seen in a wide variety of cancers[156]. p53 protein acts as a key cell cycle regulator therefore, mutations which inactivate p53 cause large increases in susceptibility to cancer. Li-Fraumeni syndrome is a rare inherited syndrome which stems from an inherited mutation in the *p53* gene, patients with this syndrome develop multiple cancers throughout their lifetime and at an earlier age than normal[157]. Missense mutations in *p53* often lead to more

stable versions of p53 protein than wild type leading to build up of p53 in the cell[156]. It is theorised that this increases the chances of it being degraded and presented by MHC1 molecules, or cell death leads to p53 entering the blood steam and lymphatic system where appropriate B cells can detect it[156]. Interestingly the majority of antibodies produced against p53 bind to the terminal regions not the mutated site[145].

1.5.2 Aberrantly expressed TAA

There exists a large body of evidence that many TAA are aberrantly expressed proteins. This can stem from proteins being constituently overexpressed[158,159] or expression in unnatural tissues or at the wrong time. Sahin et al. (1995)[160] showed MAGE1 expression, usually only seen in the testis and during development, in melanoma tissue. They were also able to demonstrate AAbs against MAGE1 present in the sera of the melanoma patients. Many other groups have shown expression of cancer testis antigens and embryonic antigens expressed in tumour tissue as well as a corresponding AAb response, such as NY-ESO- 1[161] and SOX2[162].

1.5.3 Post-translationally modified TAAs.

Cell membrane protein MUC1 is often aberrantly glycosylated in cancer. MUC1 is usually highly O-glycosylated in a 20 AA repeat region. Wandall et al. (2010)[163] have shown that an aberrantly glycosylated form of MUC1, Tn/STn-Muc1, which is commonly seen in breast cancer, triggers dendritic cell presentation on its MHC class 1 molecules not seen with the highly glycosylated form. Aberrantly glycosylated forms of MUC1 have been shown to be overexpressesed in some HCC cell lines[164].

1.5.4 Autoantibody investigation methods

AAb research has been conducted on most common cancers. Multiple methods have been developed for antigen discovery. Serological analysis of expression cDNA libraries (SEREX) involves producing a recombinant antigen library from the mRNA of a specific cancer sample[165]. This library is then screened against blood samples from multiple patients to look for reactivity. This technique is limited due to the individual nature of each person's tumour proteome. SEREX

also suffers from being very labour and time extensive when constructing the cDNA library. Phage-Display overcomes this, by offering quick automated cDNA library construction[166]. Antibodies in serum then bind to the human peptide sequences on the phage coat. Bound phage can be captured and sequenced to discover the peptide sequence, this is however time consuming and costly. These techniques have outdated more traditional methods of antibody detection, such as 2D western blots (Also known as serological proteomics-based approach, SERPA)[167]. Protein microarrays have made for quick turnaround time, as pre-printed arrays can be purchased direct from manufacturers and used instantly, producing quick results but are limited by cost and can suffer from issues created by background noise. Antibody capture enzyme-linked immunosorbent assay (ELISA) produces high signal quality, with low background noise. However, ELISA suffers from long and often costly antigen preparation steps, and limited space for antigens on each plate (96 on a standard micro-well plate for ELISA vs approx. 16000 on a protein microarray). This can be partially overcome using robotic pipetting machines to increase throughput.

1.5.5 Clinical utility of Autoantibodies as biomarkers

The primary focus of cancer AAb studies is use as a biomarker for the presence of cancer or differentiation of cancerous nodules from non-cancerous nodules. Research studies have identified AAbs in the blood of patients with a wide variety of cancers including: lung, liver, breast, colorectal, esophageal, gastric, ovarian, thyroid and prostate. Whilst most of these studies demonstrate a use for AAbs on their own or in panels of biomarkers, follow through to a clinical setting is rare. Two examples of tests that have progressed are EarlyCDT-Lung[168] and PAULA's test[169] both of which are used in lung cancer assessment. EarlyCDT-Lung measures levels of seven AAbs and has a sensitivity of 41% and specificity of 93%. PAULA's test measure circulating CA-125, CEA and CYFRA 21-1 proteins plus NY-ESO-1 AAb levels and has a sensitivity of 75% and specificity of 80%. Both tests have found clinical utility in the risk stratification of high-risk patients with each test fulling a different niche based around its performance

characteristics. The high specificity of EarlyCDT-Lung lends its self to risk stratification of nodules detected by CT screening. The higher sensitivity but lower specificity of PAULA's test makes it more suitable for prioritising patients that should undergo lung cancer screening via CT.

1.6 Autoantibodies in HCC

A literature search for articles which give enough information to calculate specificity and sensitivity values of AAbs in HCC returned 20 published papers (see section 4.1.1). However, agreement on which TAAs produce an AAb response in HCC is varied; this reflects the heterogeneity of tumours. AAbs to some antigens are described multiple times such as p53, p62, C-MYC and Ral-A. No single AAb shows a high enough performance to perform as a standalone test. Oncimmune Holdings PLC (Oncimmune), among others, have shown that a panel of multiple TAAs can be useful in detecting liver cancer[170–172]. This requires screening of large numbers of TAAs to find those with high sensitivity and specificity values that identify independent groups of patients. Combining two high performing TAAs may not be as beneficial as combining one high performance and one medium performance antigen, depending on patient overlap. In heterogeneous diseases, such as HCC, each subtype may have its own set of unique biomarkers. Therefore, a biomarker that is excellent for detecting a particular subtype might only have low sensitivity for the disease overall as its sensitivity is capped by the prevalence of that subtype[173,174].

Chapter 2. Aims of the project

Previous pilot studies, in collaboration with Oncimmune, identified a panel of seven antigens useful for the screening and detection of early stage HCC[170]. However, the performance requires improvement for clinical utility. Previous work by Oncimmune suggests that expansion of a 7 antigen panel for lung cancer, EarlyCDT-lung[168], to 14 antigens can significantly improve accuracy (unpublished data).

This project aims to identify, produce and test antigens, with literature evidence for involvement in HCC, with the view to improve the current panel, especially for detection of patients with early stage disease. Antigens will be identified from two strands of the literature; TAAs shown to produce an AAb response which have previously been untested by Oncimmune and proteins with strong links to HCC which have not been tested for AAb response in HCC as of yet. Top performers will be analysed with the aim to create an expanded panel of AAbs with increased performance over the current seven AAb panel. This new panel will undergo training assays (assays performed from which the data will be used to select the panel and set cut-offs) and after identification of suitable patient cohorts, the performance of the new panel will be clinically validated. Current sample collection routes point towards collection of HCC from a background of HCV, which would show clinical utility of the panel in at least one high risk group of patients.

This project also aims to obtain sufficient serial samples from high risk liver disease patients, many of which will go on to develop HCC. This would allow monitoring of AAb levels prior to clinical diagnosis of HCC and investigation of whether developing personal cut-offs for each patient's AAb profile would enable earlier detection of HCC. A similar method has been successfully implemented for the monitoring of patients with high risk of ovarian cancer, using an algorithm named ROCA[175]. Extensive data collection on high risk HBV and HCV patients will also allow for analysis of effect on AAb response in these two separate patient groups. This individual patient tracking may also be useful in monitoring for reoccurrence after treatment.

In summary this project aims to:

1. Identify proteins with literature evidence of involvement of HCC development which may generate measurable levels of AAbs in patients with HCC
2. Clone, express and purify these proteins for use in ELISA to measure AAb responses in cohorts containing samples from HCC patients, patients with non-cancerous chronic liver disease and healthy controls
3. Integrate this data with previous studies by Oncimmune to improve upon the accuracy and early stage sensitivity of the initial panel
4. Determine a newly improved panel, with greater sensitivity for early stage disease than AFP, through training studies and clinically validate this in an independent cohort
5. Collect serial samples from patients who develop HCC to show how AAb levels change prior to, and during, HCC development

Chapter 3. High-throughput antigen production

3.1 Introduction

To identify new antigens for the panel, an extensive review of the literature was performed and from this a scoring system was developed to rank potential targets. The evidence for new antigens has been split into two; antigens that have been shown to elicit an autoimmune response in relation to HCC, and proteins that are highly associated with the development of HCC. These two targets required separate ranking systems to identify priority targets. These ranking systems are not designed to give a definitive hierarchy of antigens, only to provide help in choosing targets from a long list of candidates.

3.2 Methods

3.2.1 Identification of potential biomarkers from the literature

3.2.1.1 *Antigens with existing evidence of an immune response*

A search on Web of science, Medline and Embase was performed using the following search term; (((Autoantibod*[Title/Abstract]) AND Hepatocellular carcinoma[Title/Abstract]) AND English[Language]). For the ranking of antigens shown to elicit an immune response, only papers where the specificity and sensitivity of the marker was provided or able to be calculated were included. This allowed the assignment of a simple score, using the Youden's J statistic[176]:

$$J = \frac{true\ positives}{true\ positives + false\ negatives} + \frac{(true\ negatives)}{true\ negatives + false\ positives} - 1$$

A score of 0 means the marker has no use as it gives the same proportion of positive results for groups with and without the disease, and a score of 1 would be a perfect marker. A J score of 0.05-0.1 was given a value of 1 and a score >0.1 was given a value of 2. The J score was then combined with a number of factors revolving around the ability to detect early stage disease and the sample set used (Table 2). One point was assigned if the marker was able to identify early stage disease (Stage I or II). A point was also assigned if the authors used NCCLD controls (chronic hepatitis, FLD, ALD, cirrhosis etc.) as well as healthy controls. Finally, one point was assigned if the total cohort size

assayed was greater than 5 per marker assayed (e.g. sample size of 50 if 10 markers were tested) and 2 points if greater than 10, this was used to help ensure sufficient sample size to give a reliable result. The points assigned were totalled and the antigens ranked by score. Any equal scorers were positioned by their J score (highest J score ranked top). Any antigens previously cloned by Oncimmune were automatically included and any antigens which had previously been tested as part of the previous HCC pilot study[177] were automatically excluded as they had already been sufficiently investigated.

Table 2 - Score assigned for each factor in the literature review. Final score for each potential marker was the sum of each factor

Factor	Score assigned
Youden's score	
<0.05	0
0.05-0.1	1
>0.1	2
Sensitivity for early stage disease?	
no	0
yes	1
Chronic liver disease controls used?	
no	0
yes	1
Total cohort size per marker tested	
<5	0
5-10	1
>10	2

3.2.1.2 *Antigens related to HCC*

A different approach had to be taken with HCC associated antigens with no AAb link. There is an abundance of papers linking antigens with HCC, whose review would have required a full project in itself. Instead, a list of antigens was compiled from available published literature reviews and meta analyses in the last 10 years. These reviews were found by using the search terms on NCBI (((marker[Title/Abstract] OR biomarker[Title/Abstract])) AND (HCC[Title/Abstract] OR "primary liver cancer"[Title/Abstract])) AND (screening[Title/Abstract] OR diagnos*[Title/Abstract] OR prognos*[Title/Abstract]) with the filters review, past 10 years and English

language enabled. Abstracts were reviewed for relevance and those antigens identified as having a relationship with HCC, were recorded. For each antigen from this list, a search was performed on NCBI, for "*antigen name* and HCC". The antigens were then listed by number of papers published under the assumption that a higher volume of papers published on a given antigen would mean a stronger link with HCC. As this is a far from accurate method, further review was needed of each antigen before selection. The target antigens were selected through their association with early stage HCC and overexpression or expression of a variant in association with HCC. This was performed in descending order until 10 were selected. An additional 14 antigens were also included as Oncimmune had already cloned them for other projects or would overlap with other projects.

3.2.2 **Antigen Cloning**

In total 34 antigens required cloning; AIF1, AKR1B10, APOA1, BCL2, CDKN1B, CHGA, CPS1, DDX3X, EEF2, F2, FUCA1, GLUL, GOLM1, HGF, HNRNPA2, HNRNPL, HSPA2, HSPA4, HSPD1, IL6, IL8, KRT23, LMNB1, MDM2, MMP9, NPM1, PEBP1, PRL, RGN, SALL4, SPP1, TGFB1, XRCC5, YWHAZ. See Table 12 in section 3.3.1.3 for full details on all proteins selected.

3.2.2.1 ***Antigen isoform/variant***

As multiple isoforms/variants exist for many of the antigens, the canonical isoform or variant (as listed on www.uniprot.org) was chosen, unless a specific isoform or variant was described in the paper in which it was identified. An antibody epitope prediction tool (http://tools.immuneepitope.org/bcell/)[178] was used to select which of Oncimmune's cloning vectors was utilised for cloning; if an antigen had a predicted epitope towards the N or C terminal, the vector which placed the tags at the opposite end was used (both cloning vectors have a histidine purification tag and an assay control tag). In the case of no epitope at either terminal, the end furthest from an epitope was chosen. If a strong possibility of an epitope appeared at both terminals, the antigen was split into fragments as well as being cloned in full. This was performed to reduce any possibility of the tags altering the folding of epitopes or blocking

the binding of antibodies. Any proteins with a transmembrane domain had the domain removed and the extracellular and intracellular domains cloned into separate vectors. Knowledge of prior expression issues with large proteins meant any antigen greater than 100kDa was also split into multiple fragments. The default place was to split these in half, unless this interfered with predicted or known epitopes, in which case a region with low epitope prediction was chosen as the split point.

3.2.2.2 *Vector*

Two vectors developed by Oncimmune, NLIC and CLIC, were used for cloning[179]. The names indicate the location at the 5' or 3' of the cloning region for the DNA encoding N or C terminal AA purification and assay control tags in ligation independent cloning vectors. These are modified from the pET45b and pET21b vectors (MERCK), respectively. The vectors were modified to incorporate an *E. coli BirA* biotin-protein ligase enzyme recognition sequence (BirA). The NLIC vector contains a six-histidine tag sequence and a *BirA* sequence upstream of the cloning site, whereas these are downstream of the cloning site for the CLIC vector (Figure 7).

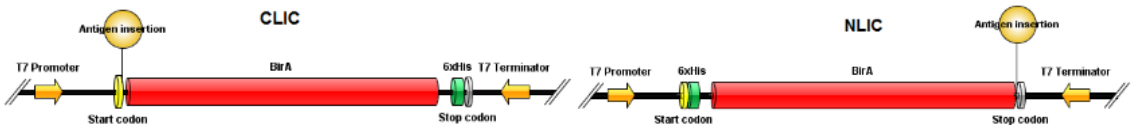


Figure 7 – Expression region of the NLIC and CLIC vectors used. Antigen insertion relates to the linearization site of the vector, where the antigen DNA is inserted

3.2.2.3 *Vector production*

JM109 *E. coli* cells (Promega) containing the CLIC or NLIC cloning vector, stored with glycerol, were removed from -80˚C storage and plated onto LB Agar containing 50µg/ml Carbenicillin (CB). A single colony was picked to inoculate 200ml of LB Broth containing appropriate antibiotics and incubated overnight at 37˚C with shaking at 200rpm. Cell pellets were formed by centrifugation at 15,000xg for 10 mins. Plasmid was isolated from the cell pellets using the Nucleobond Xtra Maxi kit (Macherey-Nagel) following the manufacturer's protocol. Vector DNA was subsequently linearised using *Stu*I and *Spe*I

restriction enzymes (NEB) for NLIC and CLIC respectively, following the manufacturer's protocol. To remove 5' overhangs, linearised vectors were subsequently treated with 2µl (20 units) mungbean enzyme (NEB) for one hour at 37˚C and subsequently inactivated by addition of 0.1% SDS. Vector was purified by gel extraction from a 0.6% agarose gel using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) following the manufacturer's protocol.

3.2.2.4 *Insert amplification*

Template cDNA was purchased from the I.M.A.G.E consortium through Source Bioscience, except for MDM2. MDM2 cDNA was purchased from Sino Biological INC. A single colony for each construct was picked to inoculate 5ml LB Broth containing the relevant antibiotic for the plasmid and grown for 16h overnight at 37˚C with shaking at 200rpm. Cells were pelleted by centrifugation (13,000g, 5 minutes), and plasmid DNA was subsequently isolated from the cell pellets using the Macherey-Nagel NucleoSpin plasmid kit (following manufacturer's instructions). Primers were picked manually, aiming for a predicted melting temperature of 68˚C, and a GC content of between 40-60%. As the whole protein was to be cloned, flexibility of primer design was limited. A sequence homologous to the 5' and 3' regions of the vector cloning site were added to the forward and reverse primers respectively (Table 3). All reactions used Q5 hot start master mix (NEB), a high-fidelity polymerase, to minimise mutations (Table 4). Firstly, amplification was attempted following the manufacturers protocol (Table 5). For any failed amplifications, a touchdown PCR protocol was used (Table 6). To determine success/failure of the reaction, 5µl of DNA from was analysed for the correct band size via electrophoresis using a 1% agarose gel (110v, 30 minutes). Unamplified vector DNA in the successful reactions was digested with 500 units DPN1 (NEB) at 37˚C for 60 mins then inactivated at 80˚C for 20 mins. DPN1 treated inserts were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) following the manufacturer's protocol.

Table 3 - Primer design

CLIC	Forward: 5'-TTTAAGAAGGAGATATACATG+14-21nt of antigen
	Reverse: 5'- AGAAGAAGAAGACGA+14-21nt of antigen
NLIC	Forward: 5'- AGTAGTGGTAGGCCG+14-21nt of antigen
	Reverse: 5'- TAGGTTAATTAAGGCTA+14-21nt of antigen

Table 4 – PCR reagents for insert amplification

Reagent	**Amount**
DNA template	10ng
Q5 Hot start master mix	25µl
Forward primer	10µM
Reverse primer	10µM
Nuclease free water	Make up to 50µl

Table 5 – Standard PCR conditions

Temperature (˚C)	**Time (seconds)**	**Cycles**
98	180	X1
98	10	X35
68	20	
72	15	
72	120	X1
4	Hold	X1

Table 6 – Touchdown PCR conditions

Temperature (˚C)	**Time (seconds)**	**Cycles**
98	180	X1
98	10	X30
70-55 (dropping 0.5˚C per cycle)	20	
72	15	
98	10	X15
55	20	
72	15	
72	120	X1
4	Hold	X1

3.2.2.5 *Gibson cloning*

Cloning was accomplished using Gibson cloning master mix (NEB)with a 3:1 molar ratio of insert to vector following the manufacturers protocol, and subsequently 2µl of the resulting cloning mixture transformed into 50µl of NEB 5-alpha *E. coli* cells via heat shock (Table 7) and plated on LB Agar + 50µg/ml CB (37˚C, 16 hours). Primers corresponding to the T7 promotor and terminator, flanking the cloning site and tags these two sites were used for colony PCR to check for insert incorporation. A sterile pipette tip was used to pick colonies and then inoculate the PCR mixture (Table 8 & Table 9), and subsequently placed in a 50ml falcon tube containing 5ml LB Broth + 50µg/ml CB. Success/failure was determined by analysing 5µl of each PCR reaction by gel electrophoresis (see section 3.2.2.4). Successful colonies were grown overnight and plasmid DNA isolated as described in section 3.2.2.4. If colonies did not contain the insert, further colonies were picked until a successful transformant was found. If no successful colonies were identified after three attempts, the transformation was repeated with the additional step of treating the reaction with DPN1 and purified as described in section 3.2.2.4.

Table 7 - Heat shock protocol

Step	**Time**
Ice incubation	30 mins
Heat shock (42˚C)	30 seconds
Ice incubation	5 minutes
Addition of 950µl SOC media and incubation at 37˚C	90 mins

Table 8 - Colony PCR reagents

Reagent	**Amount**
Q5 Hot start master mix	12.5µl
T7 Forward primer (AATACGACTCACTATAGGG)	10µM
T7 Reverse primer (TTATGCTAGTTATTGCTCAG)	10µM
Nuclease free water	Make up to 25µl

Table 9 - Colony PCR conditions

Temperature (˚C)	Time (seconds)	Cycles
95	300	X1
95	60	X30
54	90	
72	60	
72	300	X1
4	Hold	X1

3.2.2.6 *Sequencing*

Successful clones were sent for sequencing to Source BioScience. Data files were analysed using Serial Cloner (Serial Basics) and Sequencher (Gene codes Corp.) to align sequences to the reference sequence. A decision was made on a case by case basis to accept or reject any mutations. All frameshift mutations were rejected and another successful transformant sent for sequencing.

3.2.3 High-throughput protein expression

3.2.3.1 *Transformation*

BL21(DE3)RIPL *E. coli* cells (Agilent) were transformed with 2µl of antigen containing vector (Table 10). Transformed cells were plated and grown overnight on LB agar containing 50µg/ml CB, 30µg/ml streptomycin (STREP) and 34µg/ml chloramphenicol (CM, 37˚C, 16 hours).

Table 10 - BL21(DE3) RIPL transformation protocol

Step	Time
Ice incubation	30 minutes
Heat shock (42˚C)	20 seconds
Ice incubation	2 minutes
Addition of 180µl SOC media and incubation at 37˚C	90 minutes

3.2.3.2 *Starter Cultures*

Starter cultures were used to ensure sufficient inoculation volume for expression. LB Broth (10ml) supplemented with 50µg/ml CB, 34µg/ml CM and 1% Glucose, was inoculated with a single colony and incubated at 37˚C for 16 hours overnight with shaking at 200rpm.

3.2.3.3 *Expression cultures*

60ml of auto-inducing phosphate buffered Turbo Broth Overnight expression medium (TBOnex-P) containing 50µg/ml CB was inoculated with 1/20th volume of starter culture and expression induced after two hours incubation at 37˚C. The 60ml cultures were then incubated at 25˚C for 18 hours with shaking at 200rpm. After incubation cells were harvested by centrifugation (14,000g, 4˚C, 30 minutes), and the supernatant discarded. Cell pellets were frozen at -80˚C until purification.

3.2.4 **Full scale expression**

See section 5.2.1

3.2.5 **HTP Purification**

3.2.5.1 *Soluble Lysis*

Cell pellets were removed from -80˚C storage and brought to room temperature. Soluble lysis was accomplished by resuspending cell pellets in 250µl soluble lysis buffer. After 20 minutes rolling incubation at room temperature, tubes were centrifuged (14,000g, 30 minutes), the supernatant collected and transferred to another tube for storage at 4˚C. This process was repeated with ½ the initial volume of soluble lysis buffer and the supernatant was combined with the supernatant from the first incubation and labelled as the soluble lysate. The remaining pellet is the insoluble material including the Inclusion body (IB) which is primarily made up of aggregated misfolded protein[180].

3.2.5.2 *Inclusion body lysis*

The Insoluble pellet was resuspended in IB lysis, and left rolling for 16-20 hours overnight. After centrifugation (15,000g, 30 minutes), the supernatant was stored and labelled as the IB lysate, the remaining completely insoluble material was discarded. The IB lysate was stored at 4˚C until use the same day.

3.2.5.3 *Denaturing Purification*

Each well of a 24 well 25µm filter plate was loaded with 1ml of nickel sepharose resin FF crude Ni-NTA (GE Healthcare) and equilibrated in either guanidine

wash buffer or soluble binding buffer, for insoluble or soluble purification respectively. A vacuum manifold was used for all liquid steps except for elution. Lysate was added to the wells and pulled through. The resin was then washed with 5ml guanidine wash buffer, followed by 5ml urea wash buffer. Denatured protein was refolded in 5ml urea refolding buffer then conditioned in 5ml of conditioning buffer. Proteins were eluted in 5ml of elution buffer by centrifugation (2,000g, 1 minute) and collected in a 10ml 24 well plate.

3.2.5.4 *Purification analysis*

3.2.5.4.1 Coomassie stained SDS-PAGE

Purified proteins were denatured and reduced using NuPage LDS sample buffer (Novex) with 50mM DTT and subsequently run on 4-20% gradient Precast RunBlue SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) gels (Expedeon) for 45 minutes at 170v. Total protein was detected by staining gels in 50ml InstantBlue Coomassie stain (Expedeon) for one hour and subsequently de-stained using dH_20 for one hour to reduce background staining. Novex sharp protein standards were used to identify molecular weight of the present bands with presence of a band at the predicted molecular weight inferring successful purification.

3.2.5.4.2 Bradford assay

Quantification was primarily accomplished via a Bradford assay using the Bio-Rad Coomassie reagent. Protein samples were added to a 96 well plate in triplicates, along with a range of BSA standards from 0µg/ml-500µg/ml. The Bio-Rad protein assay reagent was added to each well and the OD measured at 595nm range using a Tecan infinite plate reader. Concentrations were assessed by comparison with the standard curve formed from the BSA standards.

3.2.5.4.3 Anti-HIS ELISA protein quantification

Any wells showing poor concentrations (50-100µg/ml) were checked for performance in an anti-HIS tag ELISA assay against standards produced to a higher purity. Proteins were diluted and coated as described in section 6.2.2, assuming a concentration of 100µg/ml (minimum required concentration for

the AAb ELISA). Three proteins of known concentration showing high, medium and low anti-HIS reactivity were used as comparisons. ELISA assay was then performed as in section 4.2.4, using 50µl of mouse anti-HIS antibody (Abcam, Ab15149) in place of serum samples and a HRP conjugated rabbit anti-mouse IgG (Dako, P0260) as the detection antibody. Dilution response curves were plotted for the protein of interest and the control proteins and compared at 160nM and 50nM titrations.

3.2.6 Full scale purification

See section 5.2.2

3.3 Results

3.3.1 Antigen selection

3.3.1.1 *Antigens with existing evidence of an immune response*

The literature review returned 63 unique AAb markers from 20 papers (Figure 8). Of these, seven antigens were included due to either cross over with another Oncimmune project, so were deemed efficient to produce, or having been previously cloned and expressed for other projects and 15 antigens were automatically excluded having been tested previously by Oncimmune. From the top 20 remaining antigens, 13 showed enough evidence to warrant testing, through their association with liver processes and/or appearing multiple times in the literature review. Rankings of the AAb markers are shown in Table 11. Scores ranged from 1 to 6 with an average score of 3.73. The antigens selected based in their score alone ranged from scores of 4 to 6 with an average score of 5. Antigens included due to overlap with other projects had scores ranging from 3 to 4, with an average score of 3.67.

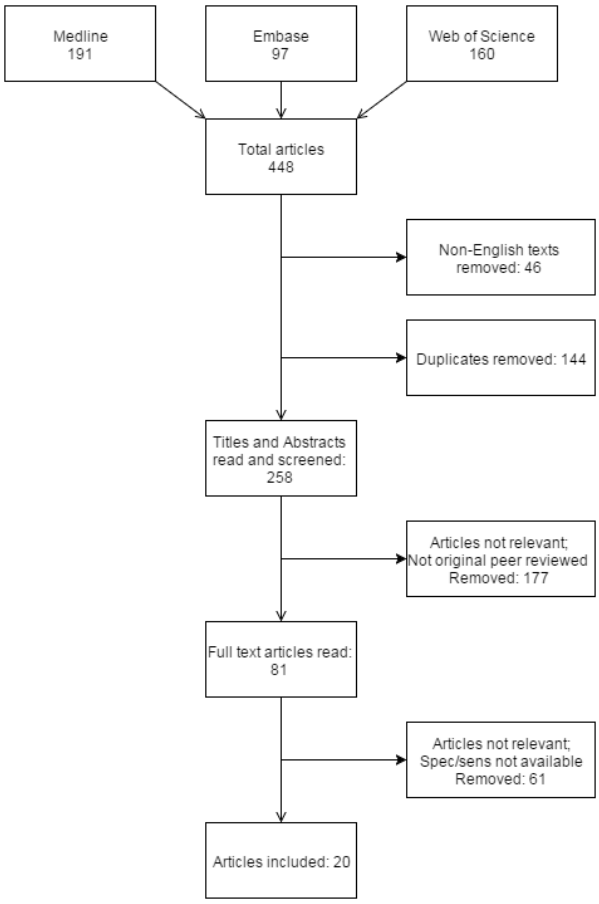


Figure 8 - Selection of appropriate AAb papers

Table 11 – Results of the AAb scorring system. Cells highlighted in green represent antigens chosen for use in this thesis based on their score. Cells highlighted in yellow represent antigens excluded from this thesis due to use in our groups previous work[177]. Cells highlighted blue represent antigens selected due to overlap with other projects. Grey cells represent antigens not selected for use in this thesis based on their score.

Antigen	Gene name	Youden's J Statistic	performance score	Early stage sensitivity	NCCLD controls used	Cohort/Marker tested	Total Score	Ref.
HNRNPL	HNRNPL	0.662	2	2	1	1	6	[181]
Dead Box 3	DDX3X	0.548	2	2	1	1	6	[182]
eEF2	EEF2	0.535	2	2	1	1	6	[182]
Ku86	XRCC5	0.511	2	2	1	1	6	[183]
Ku86	XRCC5	0.507	2	2	1	1	6	[184]
AIF	AIF1	0.414	2	2	1	1	6	[182]
HNRNPA2	HNRNPA2B1	0.409	2	2	1	1	6	[182]
Prostatic binding protein	PEBP1	0.326	2	2	1	1	6	[182]
GRP78	HSPA5	0.355	2	1	1	1	5	[185]
Survivin	BIRC5	0.204	2	1	1	1	5	[172]
RalA	RALA	0.194	2	1	1	1	5	[186]
Caper Alpha/HCC1	RBM39	0.191	2	1	1	1	5	[172]
RalA	RALA	0.188	2	1	1	1	5	[171]
RalA	RALA	0.164	2	1	1	1	5	[172]
Koc	IGF2BP3	0.159	2	1	1	1	5	[171]
p62	IGF2BP2	0.155	2	1	1	1	5	[171]
p62	IGF2BP2	0.151	2	1	1	1	5	[172]
Imp1	IGF2BP1	0.147	2	1	1	1	5	[171]
p53	TP53	0.147	2	1	1	1	5	[171]
p16	CDKN2A	0.146	2	1	1	1	5	[171]
MDM2	MDM2	0.138	2	1	1	1	5	[172]

14-3-3 zeta	YWHAZ	0.132	2	1	1	1	5	[172]
Cyclin B1	CCNB1	0.132	2	1	1	1	5	[172]
CyclinB1	CCNB1	0.129	2	1	1	1	5	[171]
C-myc	MYC	0.11	2	1	1	1	5	[171]
Sui1	EIF1	0.11	2	1	1	1	5	[171]
c-Myc	MYC	0.105	2	1	1	1	5	[172]
p53	TP53	0.105	2	1	1	1	5	[172]
Koc	IGF2BP3	0.104	2	1	1	1	5	[172]
survivin	BIRC5	0.103	2	1	1	1	5	[171]
HSPA4	HSPA4	0.371	2	1	1	0	4	[187]
Vimentin	VIM	0.343	2	1	1	0	4	[187]
p62	IGF2BP2	0.307	2	1	1	0	4	[187]
RALA	RALA	0.292	2	1	1	0	4	[187]
HCC-22-5	RGN	0.274	2	0	1	1	4	[188]
LMNB1	LMNB1	0.271	2	1	1	0	4	[187]
DDX3	DDX3X	0.271	2	1	1	0	4	[187]
GRP78	HSPA5	0.271	2	1	1	0	4	[187]
HSP70	HSPA2	0.238	2	0	1	1	4	[189]
p53	TP53	0.238	2	0	1	1	4	[190]
p62	IGF2BP2	0.211	2	0	1	1	4	[191]
HSPD1	HSPD1	0.167	2	1	1	0	4	[187]
KRT 23	KRT 23	0.164	2	0	1	1	4	[172]
14-3-3 zeta	YWHAZ	0.156	2	0	1	1	4	[192]
p53	TP53	0.129	2	1	1	0	4	[187]
MDM2	MDM2	0.127	2	0	1	1	4	[193]

ANXA2	ANXA2	0.126	2	1	1	0	4	[187]
SND1	SND1	0.119	2	1	1	0	4	[187]
NPM1	NMP1	0.085	1	1	1	1	4	[172]
p90	KIAA1524	0.084	1	1	1	1	4	[172]
IMP1	IGF2BP1	0.078	1	1	1	1	4	[172]
Kinectin	KTN1	0.8	2	0	1	0	3	[188]
LDH-A	LDHA	0.732	2	0	1	0	3	[188]
Albumin	ALB	0.709	2	0	1	0	3	[188]
DNA binding factor	Various	0.686	2	0	1	0	3	[188]
C3VS (5')	UACA	0.664	2	0	1	0	3	[188]
P14.5	p14.5	0.6	2	0	1	0	3	[188]
Calreticulin	CalR	0.523	2	0	1	0	3	[194]
β-Tubulin	TUBB	0.368	2	0	1	0	3	[194]
Nucleoside Diphosphate Kinase 1	NME1	0.294	2	0	1	0	3	[194]
AHSG	AHSG	0.278	2	0	0	1	3	[195]
CENPF	CENPF	0.243	2	1	0	0	3	[187]
FTL	FTL	0.215	2	0	0	1	3	[195]
F1-ATP synthase β subunit	ATP5A1	0.203	2	0	1	0	3	[194]
Cytokeratin 8	CK8	0.203	2	0	1	0	3	[194]
Creatine kinase B	CKB	0.197	2	0	1	0	3	[194]
CK18	CK18	0.133	2	0	1	0	3	[194]
NY-ESO-1	CTAG1B	0.122	2	0	0	1	3	[196]
Transferrin	TF	0.074	1	1	1	0	3	[187]
HSP90AA1	HSP90AA1	0.069	1	1	1	0	3	[187]

RAB4B	RAB4B	0.065	1	1	1	0	3	[187]
p16	CDKN2A	0.012	0	1	1	1	3	[172]
p53	TP53	0.036	0	0	1	1	2	[197]
PTGS2	PTGS2	0.026	0	1	1	0	2	[187]
Koc	IGF2BP3	0.025	0	0	1	1	2	[197]
c-myc	MYC	0.023	0	0	1	1	2	[197]
GLUD1	GLUD1	0.021	0	1	1	0	2	[187]
MAGEA2	MAGEA2	0.021	0	1	1	0	2	[187]
PRDX1	PRDX1	0.021	0	1	1	0	2	[187]
EIF3I	EIF3I	0.018	0	1	1	0	2	[187]
LGTN	LGTN	0.014	0	1	1	0	2	[187]
Imp1	IGF2BP1	0.013	0	0	1	1	2	[197]
survivin	BIRC5	0.012	0	0	1	1	2	[197]
p62	IGF2BP2	0.012	0	0	1	1	2	[197]
ANXA5	ANXA5	0.006	0	1	1	0	2	[187]
FN1	FN1	0.006	0	1	1	0	2	[187]
HNRNPA2B1	HNRNPA2B1	0.004	0	1	1	0	2	[187]
PGK1	PGK1	0.001	0	1	1	0	2	[187]
EIF3L	EIF3L	-0.019	0	1	1	0	2	[187]
HSP60	HSPD1	-0.038	0	0	1	1	2	[189]
Mage-C2	MAGEC2	0.0118	0	0	1	0	1	[198]
Hsp60	HSPD1	-0.029	0	0	1	0	1	[194]

3.3.1.2 *Antigens related to HCC*

The list of antigens associated with HCC compiled from literature reviews and meta-analyses over the past ten years totalled 63. Of these, 12 antigens were included due to either cross over with another Oncimmune project, so were deemed efficient to produce, or having been previously cloned and expressed for other projects and 10 antigens were automatically excluded having been tested previously by Oncimmune. After excluding antigens based on lack of evidence of involvement in early HCC, and unavailability of template DNA, 11 antigens were chosen.

3.3.1.3 *Antigens to be cloned*

In total this led to 43 antigens being chosen for investigation in this project, and 34 of these needing to be cloned. These antigens are listed in Table 12 along with their link to HCC.

Table 12 - Selected antigens and their link to HCC. For the rest of this thesis all antigens will be referred to by their short protein name.

Gene name	Protein name	Short protein name	Involvement in HCC	Reference
AIF1	Allograft inflammatory factor 1	AIF1	No role in hepatocarcinogenesis, however AAbs have been seen in HCC patients. Plays a role in the anti-inflammatory response to vascular damage and regulation of macrophages.	[199]
AKR1B10	Aldo-keto reductase family 1 member B10	AKR1B10	Enzyme capable of reducing aliphatic and aromatic aldehydes. Overexpressed in HCC. Correlated with early stage disease, overexpression is linked with a better prognosis than those without overexpression.	[200]
ANXA2	Annexin A2	ANXA2	Over expression of Annexin-A2 has been linked to invasion and migration in HCC. High levels are often seen in the serum of HCC patients.	[201]
APOA1	Apolipoprotein A-I	APOA1	The major protein component of HDL. Participates in the reverse transport of cholesterol from tissues to the liver. No direct carcinogenic activity identified but overexpressed in many HCCs.	[202]
BCL2	Apoptosis regulator Bcl-2	BCL2	A mitochondrial outer membrane protein involved in suppressing apoptosis. Has been shown to be overexpressed in several cancers including HCC. Also linked with chemotherapy resistance.	[203]
CD44	CD44 antigen	CD44	CD44 participates in a wide variety of cellular functions including lymphocyte activation, recirculation and homing, haematopoiesis, and tumour metastasis. CD44 is thought to be a marker of cancer stem cells. It is overexpressed in some HCCs.	[204]
CDKN1B	Cyclin-dependent kinase inhibitor 1B	CDKN1B	A cell cycle regulator often supressed in HCC. ser10 phosphorylated CDKN1B levels are often increased in HCC and correlated with poor prognosis. This modification could trigger an AAb response as seen in MUC1.	[205]
CHGA	Chromogranin-A	CHGA	A precursor molecule for paracrine and autocrine active peptides. Higher than normal levels found in chronic liver, heart and renal disease patients. Significantly higher levels found in HCC patients than NCCLD patients.	[206]

KRT18	Keratin, type I cytoskeletal 18	CK18	CK18 can be a marker of cell death in hepatocytes. Caspase associated apoptosis leads to cleaved CK18 being released into plasma.	[207]
KRT19	Keratin, type I cytoskeletal 19	CK19	HCCs which stains positive for CK19 shows a significantly worse prognosis than negative staining HCC.	[208]
CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	CPS1	The target protein of a HCC specific immunostaining antibody Hep-par-1. A liver specific enzyme that is a rate limiting factor in the urea cycle. Current results are contradictory; Hep-par-1 staining has been linked with HCC, although one group has linked suppression of CPS1 with HCC.	[209]
DDX3X	ATP-dependent RNA helicase DDX3X	DDX3X	A RNA helicase that is required for HCV replication. DDX3X AAbs have been seen in HCC patients	[210]
EEF2	Elongation factor 2	EEF2	EEF2 is crucial for protein synthesis, catalysing the translocation of elongated peptidyl-tRNA from A to P sites of the ribosome. It can also promote G2/M cell cycle progression.	[211]
EPCAM	Epithelial cell adhesion molecule	EPCAM	EPCAM positive HCC cells have been linked with poor prognosis as they appear to have more characteristics of hepatic stem cells, with increased chance of metastasis and invasion.	[212]
F2	Prothrombin	F2	DCP is the product of improperly carboxylated prothrombin, commonly occurring in HCC, thought to be due to the increased expression of prothrombin precursors in HCC. High DCP serum levels are strongly linked with presence of HCC.	[213]
FUCA1	Tissue alpha-L-fucosidase	FUCA1	FUCA1 is a lysosomal enzyme present in all mammalian cells. It is involved in the catabolism of fucose-containing glycoconjugates. It is commonly overexpressed in the serum of HCC patients, however the mechanism behind this is unknown.	[214]
GLUL	Glutamine synthetase	GLUL	Catalyses the synthesis of glutamine from glutamate and ammonia. Glutamine is a main source of energy and is involved in inhibition of apoptosis, and cell signalling. Overexpressed in the tissue and serum of some HCC patients.	[215]
GOLM1	Golgi membrane protein 1	GOLM1	GP73 is a Golgi transmembrane protein. It processes proteins synthesized in the rough endoplasmic reticulum and assists in the transport of protein cargo through the Golgi apparatus. GP73 is found at high levels in the serum of HCC patients and has been touted as a	[216]

			more specific replacement for AFP. Reasons for overexpression/its role in carcinogenesis are not well understood.	
HGF	Hepatocyte growth factor	HGF	HGF is the ligand for the proto-oncogene MET. In adults, MET expression is limited to stem and progenitor cells and is necessary for wound healing and hepatocyte regeneration. Aberrant activation of the HGF/MET pathway leads to a variety of cancers. High HGF serum levels have been observed in HCC patients.	[217]
HNRNPA2	Heterogeneous nuclear ribonucleoproteins A2	HNRNPA2	hnRNP-A2 up-regulation in HCC induces an alternative splicing switch that down-regulates a dominant-negative isoform of A-Raf, leading to activation of the Raf-MEK-ERK pathway and cellular transformation.	[218]
HNRNPL	Heterogeneous nuclear ribonucleoprotein L	HNRNPL	In HCC cell cultures, hnRNP-L knock down demonstrated that this protein sustains the growth, migration and invasion of the cancer cells. An AAb response has been shown to differentiate HCC patients from chronic HBV patients	[219]
HSPA2	Heat shock-related 70 kDa protein 2	HSPA2	HSPA2 is a member of the HSP70 family of proteins involved in protein folding and avoiding aggregation. It has been shown to be overexpressed in HCC tissue samples; overexpression has been linked with poor prognosis.	[220]
HSPA4	Heat shock 70 kDa protein 4	HSPA4	Upregulation of HSPA4 is associated with a poor outcome from early stage HBV related HCC.	[221]
HSPD1	60 kDa heat shock protein, mitochondrial	HSPD1	There are no strong links between HSPD1 and HCC however an AAb response has been shown in some HCC patients.	[195]
IL6	Interleukin-6	IL6	IL-6 is a major cause of inflammation in the liver and has been linked with facilitating HCC by activating several pathways which increase cell proliferation, protect from apoptosis and increase metastatic potential.	[222]
IL8	Interleukin-8	IL8	IL-8 is a pro-inflammatory cytokine found at high levels in the serum of HCC patients. It can also promote angiogenesis, tumorigenesis and self-renewal.	[223]

KRT23	Keratin 23	KRT23	High tissue levels of KRT23 can serve as a marker of severe steatohepatitis and potentially early HCC.	[224]
LMNB1	Lamin-B1	LMNB 1	LMNB1 is a nuclear envelope protein and possesses a transcriptional coregulatory activity having an important role in DNA replication, cellular aging, and stress responses. It has been shown to be upregulated in the tissue and sera of HCC patients.	[225]
MAGEA3	Melanoma-associated antigen 3	MAGE A3	MAGE A3 is one of the most commonly seen cancer testis antigens. It is usually only seen in the testis, however is often expressed by cancer cells. Cancer testis antigens have been proven to elicit an immune response in a variety of cancers.	[226]
MDM2	E3 ubiquitin-protein ligase Mdm2	MDM2	MDM2 is an inhibitor of p53 and functions in a feedback loop. MDM2 is often overexpressed in HCC therefore suppressing p53 activity.	[227]
MMP9	Matrix metalloproteinase-9	MMP9	MMP9 is often overexpressed in HCC and is critical for tumour invasion and progression to metastasis	[228]
MSLN	Mesothelin	MSLN	Mesothelin overexpression is seen in cholangiocarcinoma and in a variety of other cancers.	[229]
NPM1	Nucleophosmin	NPM1	NPM1 is frequently overexpressed, mutated, rearranged and deleted in human cancer. NPM1 has been shown to be overexpressed in HCC cells compared to normal hepatocytes.	[230]
PEBP1	Phosphatidylethanolamine-binding protein 1	PEBP1	PEBP1 is often down-regulated in HCC, this helps prevent drug induced apoptosis. Although it is down regulated AAbs to PEBP1 have been shown in HCC patients.	[231]
PRL	Prolactin	PRL	PRL has been shown to have both protective and carcinogenic effects in HCC. The carcinogenic effect occurs when overexpressed in combination with a prolactin receptor mutation.	[232]
RGN	Regucalcin	RGN	RGN has been demonstrated as a suppressor of cellular proliferation. RGN levels are often supressed in HCC; however, AABs targeting it have been detected.	[233]
SALL4	Sal-like protein 4	SALL4	SALL4 is a stem cell related gene which increases cell proliferation. It is overexpressed in several cancers, including HCC, and is linked with poor prognosis.	[234]

SPP1	Osteopontin	SPP1	SPP1 can induce inflammation. It has been shown in some HCCs that SPP1 is hypomethylated leading to increased SPP1 levels. Different splice variants have been shown to trigger immune reactions is various cancers.	[235]
TF	Serotransferrin	TF	Serum Transferrin levels in African-American individuals with HCV related HCC are significantly higher than in Caucasian Americans with HCC.	[236]
TGFB1	Transforming growth factor beta-1	TGFB1	TGFB1 has both tumour promoting and tumour suppressive effects in HCC, although the mechanisms behind these are not fully understood. It is thought to promote epithelial to mesenchymal transition. Overexpression can be seen in both tissue and plasma.	[237]
VIM	Vimentin	VIM	Vimentin has been demonstrated as a marker of epithelial to mesenchymal transition in several cancers. It has been shown to be upregulated in some HCCs.	[238]
XRCC5	X-ray repair cross-complementing protein 5	XRCC5	Ku86 plays a vital role in double strand break repair. It is overexpressed in the tissue of HCC patients, potentially in response to the DNA damage leading to the cancer.	[183]
YWHAZ	14-3-3 protein zeta	YWHAZ	14-3-3 proteins bind and stabilise other proteins, typically related with cell proliferation, survival, stress and apoptosis. Overexpression of 14-3-3z along with αB-crystalline can constitutively activate ERK signalling increasing risk of HCC.	[239]

3.3.2 **Cloning**

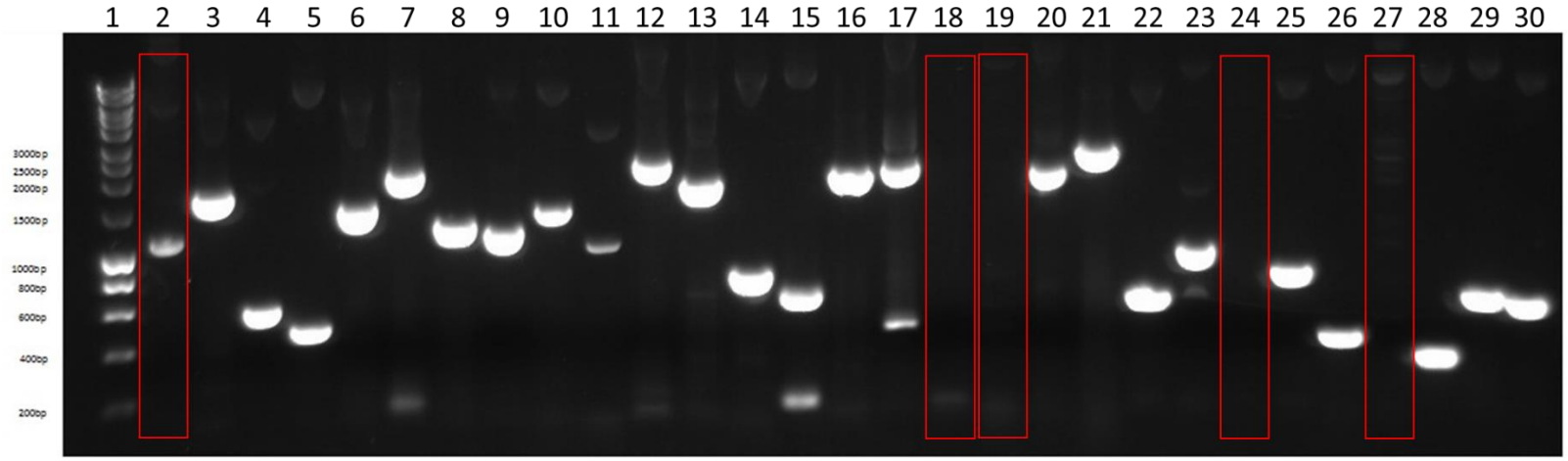
After assessing each protein for size and possible epitope binding sites, 43 constructs were designed. All constructs contained one insert, except for SALL4B, which required two inserts, as it comprises of AAs 1-384 and 822-1053 of the SALL4A protein. In total, 44 inserts needed to be amplified.

3.3.2.1 ***Insert amplification***

All insert amplifications were attempted using the manufacturer's recommended PCR conditions. Of the 44 inserts, 7 amplifications were not successful with this protocol (Figure 9) so were repeated using touchdown PCR (TDPCR) (Figure 10 &Figure 11). These are; MMP9-CLIC (wrong size band), HGF-NLIC (no band), SALL4A-CLIC (no band), KRT23-CLIC (no band), LMNB1-CLIC (no band), MDM2-CLIC (no clear band of correct size) and SALL4B-CLIC part 1 (no band). Upon first attempt with TDPCR, two amplifications (MMP9-CLIC and KRT23-CLIC) were successful (Figure 11). For the remaining five inserts, template DNA was re-prepared and standard PCR repeated. Along with these five, AKR1B10-CLIC was also amplified due to a delay in obtaining template DNA. Of the six reactions, five were successful (Figure 12), with SALL4A-CLIC the only one to fail. It was decided not to continue with SALL4A-CLIC, due to having two other variants of SALL4 which covered much of the protein sequence, and due to limitations on the number of antigens which could fit onto one HTP assay plate. More constructs were made than could fit onto one assay plate intentionally to allow for any antigens that could not be cloned.

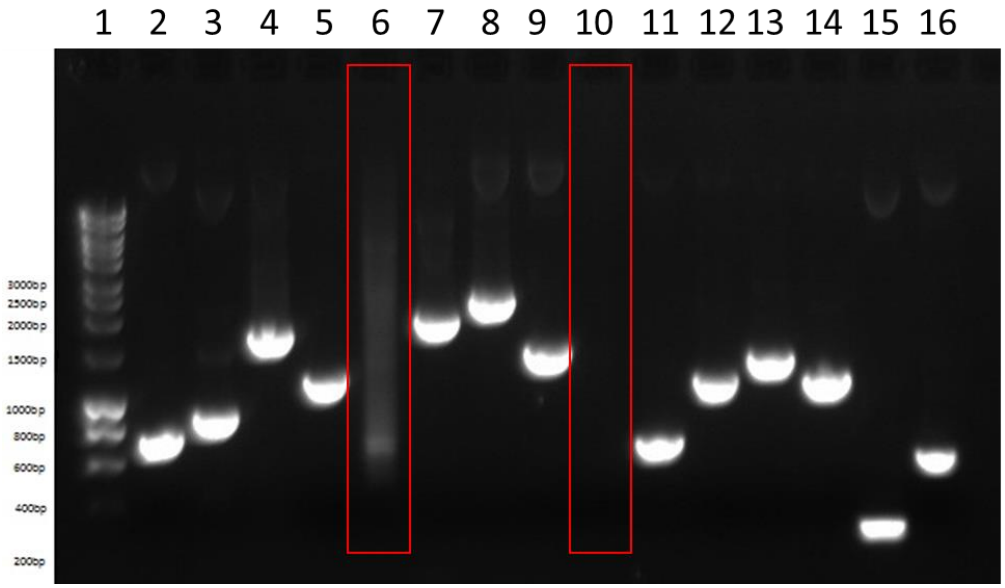
All reactions had a concentration greater than the 5ng/µl concentration needed for Gibson cloning after DPN1 treatment and purification.

Figure 9 - Insert amplification results from the first attempt with standard PCR protocol. Failed reactions are highlighted with a red box



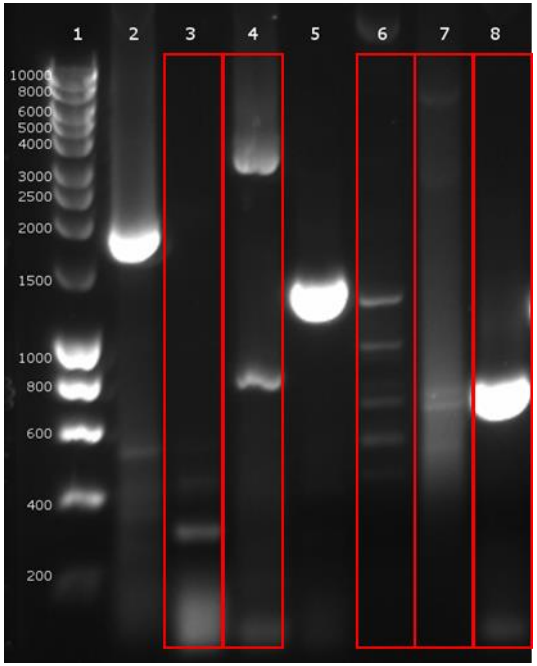
Lane	Insert	Lane	Insert	Lane	Insert
1	HyperLadder 1kb	11	SPP1-NLIC	21	EEF2-NLIC
2	MMP9-CLIC	12	XRCC5-NLIC	22	PEBP1-CLIC
3	HSPD1-CLIC	13	F2-NLIC	23	RGN-NLIC
4	HNRNPA2-CLIC	14	APOA1-CLIC	24	KRT23-CLIC
5	AIF1-NLIC	15	PRL-CLIC	25	YWHAZ-NLIC
6	FUCA1-CLIC	16	HSPA2-CLIC	26	TGFB1-NLIC
7	DDX3X-NLIC	17	CPS1-C-CLIC	27	LMNB1-CLIC
8	GOLM1-NLIC	18	HGF-NLIC	28	IL8-NLIC
9	GLUL-NLIC	19	SALL4A-CLIC	29	CDKN1B-NLIC
10	CHGA-NLIC	20	HSPA4-CLIC	30	IL6-CLIC

Figure 10- Insert amplification results from the first attempt with standard PCR protocol. Failed reactions are highlighted with a red box



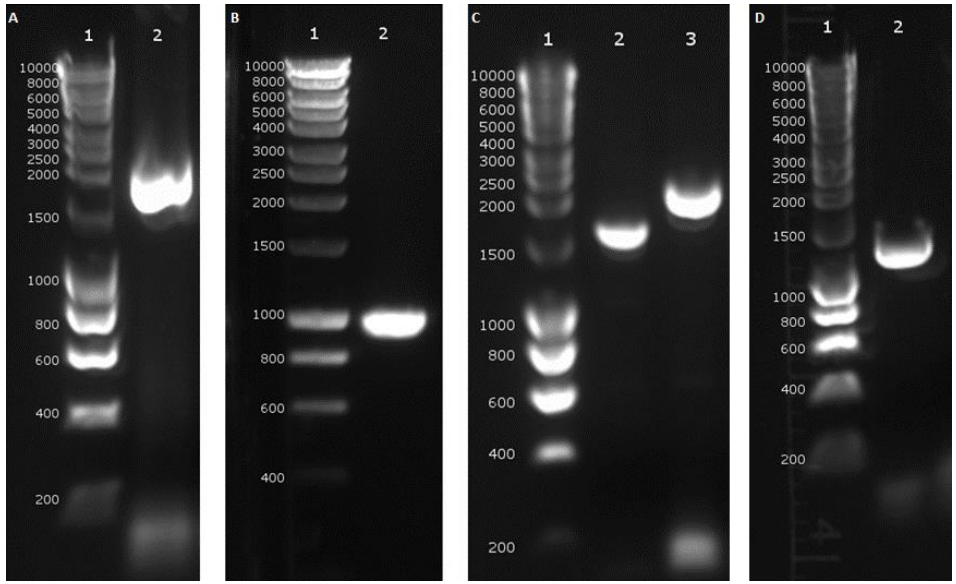
Lane	Insert	Lane	Insert
1	HyperLadder 1kb	9	CPS1-N-NLIC
2	BCL2-CLIC	10	SALL4B-CLIC part 1
3	NPM1-NLIC	11	SALL4B-CLIC part 2
4	HNRNPL-CLIC	12	SALL4-N-CLIC
5	Amplification positive control	13	EEF2-N-NLIC
6	MDM2-CLIC	14	EEF2-C-NLIC
7	DDX3X-NLIC	15	IL8-CLIC
8	CPS1-NLIC	16	CDKN1B-CLIC

Figure 11 - TDPCR insert amplification results. Failed reactions are highlighted with a red box



Lane	Insert
1	HyperLadder 1kb
2	MMP9-CLIC
3	HGF-NLIC
4	SALL4A-CLIC
5	KRT23-CLIC
6	LMNB1-CLIC
7	MDM2-CLIC
8	SALL4B-CLIC part 1

Figure 12 – Standard PCR amplification of remaining inserts using reprepared template DNA

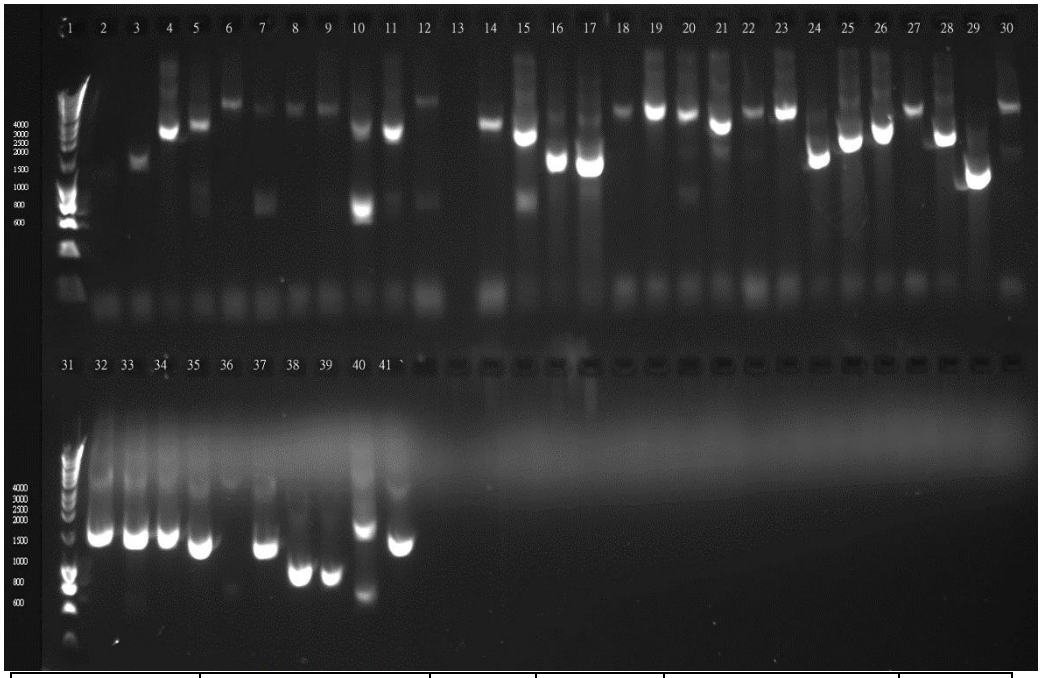


Lane	Insert
A1	HyperLadder 1kb
A2	LMNB1-CLIC
B1	HyperLadder 1kb
B2	AKR1B10-CLIC
C1	HyperLadder 1kb
C2	MDM2-CLIC
C3	HGF-NLIC
D1	HyperLadder 1kb
D2	SALL4B-CLIC part 1

3.3.2.2 *Gibson Cloning*

Gibson cloning gave colonies containing the correct size insert for 41 of the 42 remaining constructs (Example results, Figure 13). HGF-NLIC was the only construct to fail Gibson cloning. A mistake in the initial forward primer design was noticed, incorporating an extra A between the vector sequence and the antigen coding sequence. After ordering a new primer and reamplifying the insert, the cloning reaction still failed to produce any clones containing the insert, hence giving “vector only” colonies. To account for low efficiency transformation, the cloning vector was instead linearized using PCR, and subsequently treated with DPN1 to digest any circular template vector. After repeating the cloning reaction and transformation, the result was identical with “vector only” colonies. To check for successful cloning, the post cloning reaction product was analysed via gel electrophoresis. This showed a band approximately the size of the vector as well as a band the size of the vector plus the insert. This suggests the cloning reaction was successful and the problem lies elsewhere. A potential cause of this could be due to leaky expression of the HGF construct being toxic to the cells. It was decided to drop HGF as an excess of antigens were produced.

Figure 13 – Example colony PCR gel. Labels contain antigen name followed by expected size in BPs



Lane	Construct	BPs	Lane	Construct	BPs
1	HyperLadder 1		22	LMNB1-CLIC	2387
2	AIF1-NLIC	1121	23	MMP9-CLIC	2435
3	CDKN1B-CLIC	1223	24	PEBP1-CLIC	1190
4	CHGA-NLIC	1997	25	RGN-NLIC	1574
5	CPS1-N-NLIC	2180	26	SALL4-N-CLIC	1781
6	CPS1-NLIC	3290	27	SALL4B-CLIC	3844
7	CPS1-C-CLIC	2699	28	SPP1-NLIC	1622
8	DDX3X-CLIC	2615	29	TGFB1-NLIC	1016
9	DDX3X-NLIC	2663	30	XRCC5-NLIC	2873
10	EEF2-N-NLIC	2081	31	HyperLadder 1	
11	EEF2-C-NLIC	1850	32	YWHAZ-NLIC	1415
12	EEF2-NLIC	3251	33	APOA1-CLIC	1361
13	F2-NLIC	2417	34	BCL2-CLIC	1346
14	FUCA1-CLIC	2027	35	CDKN1B-CLIC	1223
15	GLUL-NLIC	1796	36	HSPA2-NLIC	2597
16	GOLM1-NLIC	1883	37	IL6-CLIC	1178
17	HNRNPA2-CLIC	1175	38	IL8-CLIC	869
18	HNRNPL-CLIC	2348	39	IL8-NLIC	917
19	HSPA4-CLIC	2600	40	NPM1-NLIC	1562
20	HSPD1-NLIC	2321	41	PRL-CLIC	1229
21	KRT23-CLIC	1895			

3.3.2.3 *Sequencing*

The remaining 41 constructs were sent for sequencing. 38 of these were fully sequence verified and matched the reference sequence. Mutations occurred in CDKN1B-NLIC, HNRNPA2-CLIC and BCL2-CLIC. CDKN1B-NLIC contained a C->T at base pair (BP) 509 changing a proline to leucine. As CDKN1B NLIC and CLIC are identical except for the location of the BirA and HIS tag, CDKN1B-NLIC was discarded. HNRNPA2-CLIC contained a triplet deletion at BP 658-660. This corresponded with a serine deletion from a linker sequence of five serines between the antigen and the BirA tag. As the linker sequence is not essential and the deletion keeps the sequence in frame, this sequence was accepted as correct. BCL2 CLIC contained a triplet insertion in linker sequence after BP 731. This inserted a serine into the linker sequence, this was accepted as correct, for the same previous reason.

3.3.3 **Expression and Purification**

3.3.3.1 *Expression*

All 40 constructs which passed sequencing plus 3 constructs previously cloned by Oncimmune were expressed via the HTP expression method. During the first two hours of incubation at 37˚C, the optical density (OD) at 600nm of each culture reached between 0.4-1 OD at which point the media automatically induces protein expression. After 18 hours of expression at 25˚C cells were harvested by centrifugation producing a cell pellet of between 3-4 grams.

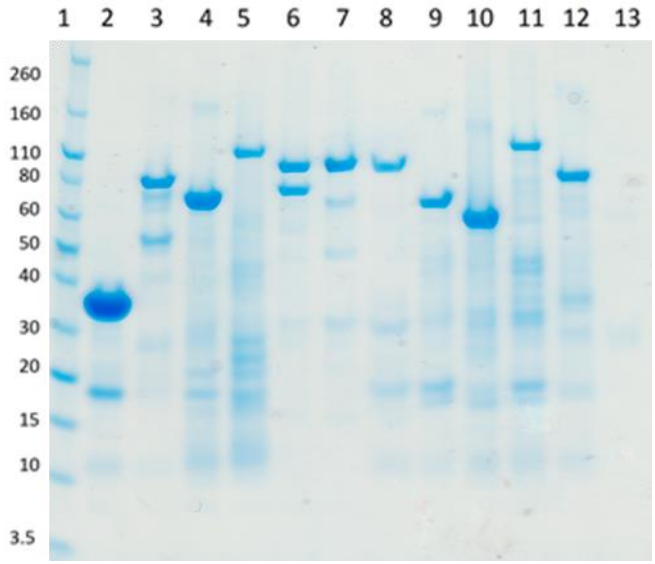
3.3.3.2 *Purification*

The harvested cell pellets underwent IB lysis and purification. Of the 43 pellets, 41 successfully produced proteins produced with varying yields ranging from 70-976µg/ml when analysed by Bradford assay (Table 133). FUCA1_C-BirA and CD44_C-BirA (cloned by Oncimmune) did not express protein at the level required for the assay.

All proteins, except for SALL4-N_C-BirA, were confirmed to have produced a protein of correct size, as confirmed by Coomassie stained SDS-PAGE (Figure 14 – SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I=

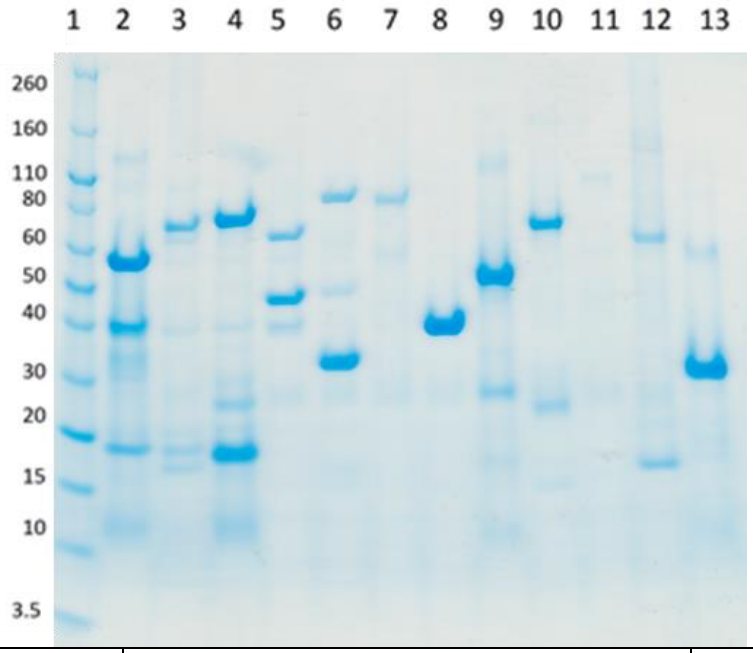
Insoluble purification-Figure 20). SALL4B_C-BirA contains all of SALL4-N_C-BirA and excess constructs were produced, SALL4-N_C-BirA was discarded and removed from the project. After multiple attempts, CD44_C-BirA and FUCA1_C-BirA could not be purified in sufficient quantity via HTP conditions to perform the discovery assay. Scaled up batches of these proteins were therefore produced (sections 3.3.3.3 and 3.3.3.4).

Five protein batches (BCL2_C-BirA, CDKN1B_C-BirA, CK19_C-BirA, MMP9_C-BirA and SALL4B_C-BirA) had concentrations of 50-100µg/ml by Bradford assay and were therefore assessed by anti-HIS ELISA. All five batches showed signal levels greater than the low control (Figure 21), so were accepted at a concentration of 100µg/ml for assay purposes.



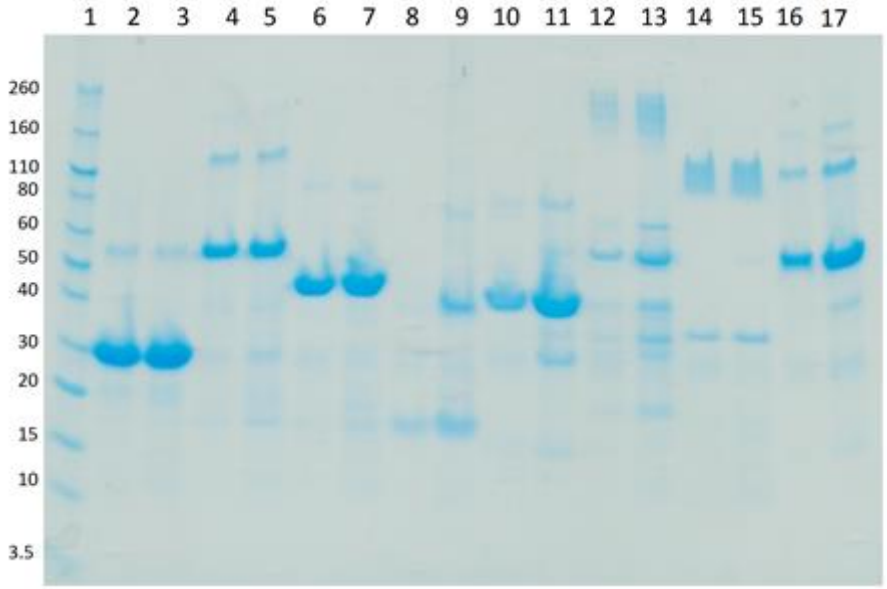
Lane	Antigen	MW (kDa)
1	Novex sharp unstained	2.3-260
2	AIF1_N-BirA I	31.6
3	CHGA_N-BirA I	63.8
4	CPS1-C_N-BirA I	70.0
5	CPS1_N-BirA I	104.9
6	CPS1_C-BirA I	89.8
7	DDX3X_C-BirA I	88.1
8	DDX3X_N-BirA I	88.0
9	EEF2-N_N-BirA I	66.8
10	EEF2-C_N-BirA I	58.2
11	EEF2_N-BirA I	110.1
12	F2_N-BirA I	80.2
13	FUCA1_C-BirA I	68.6

Figure 14 – SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification.



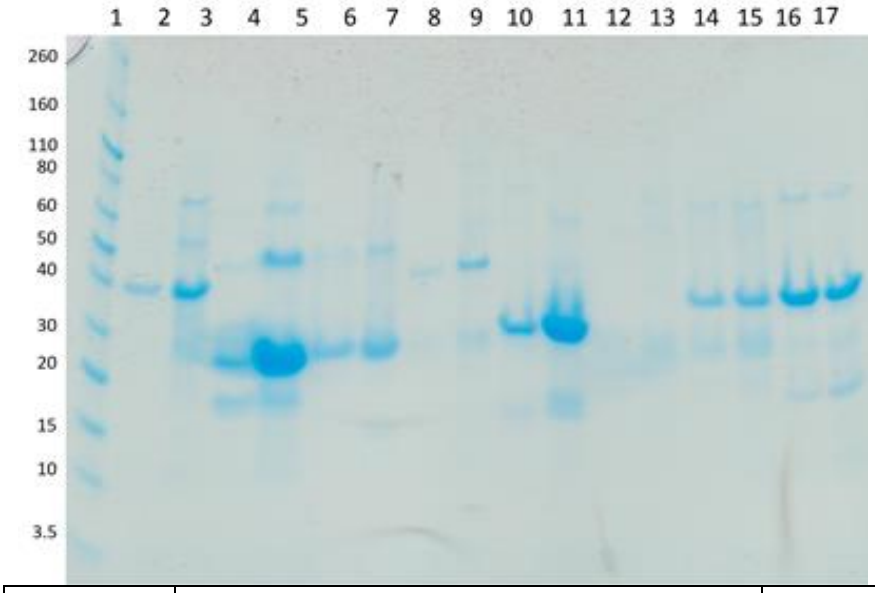
Lane	Antigen	MW (kDa)
1	Novex Sharp unstained protein standard	3.5-260
2	GLUL_N-BirA I	56.8
3	GOLM1_N-BirA I	60.2
4	HSPD1_N-BirA I	72.9
5	KRT23_C-BirA I	63.1
6	LMNB1_C-BirA I	81.3
7	MMP9_C-BirA I	81.6
8	PEBP1_C-BirA I	35.9
9	RGN_N-BirA I	48.0
10	SALL4-N_C-BirA I	55.9
11	SALL4B_C-BirA I	80.6
12	SPP1_N-BirA I	50.3
13	TGFB1_N-BirA I	27.7

Figure 15 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification



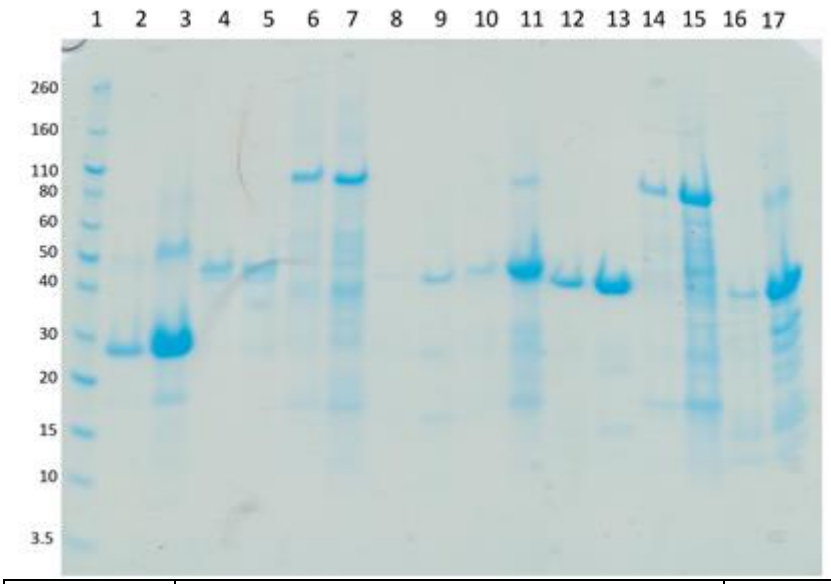
Lane	Protein	MW (kDa)
1	Novex Sharp unstained protein standard	
2	β2m_N-BirA S*	26.6
3	β2m_N-BirA I*	26.6
4	TF1_N-BirA S*	52.5
5	TF1_N-BirA I*	52.5
6	HSP27_N-BirA S*	37.6
7	HSP27_N-BirA I*	37.6
8	pIGF_N-BirA S*	34.3
9	pIGF_N-BirA I*	34.3
10	IL6_C-BirA S	35.7
11	IL6_C-BirA I	35.7
12	NPM1_N-BirA S	47.4
13	NPM1_N-BirA I	47.4
14	TTR_N-BirA S*	28.6
15	TTR_N-BirA I*	28.6
16	TF2_C-BirA S*	53.9
17	TF2_C-BirA I*	53.9

Figure 16 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project.



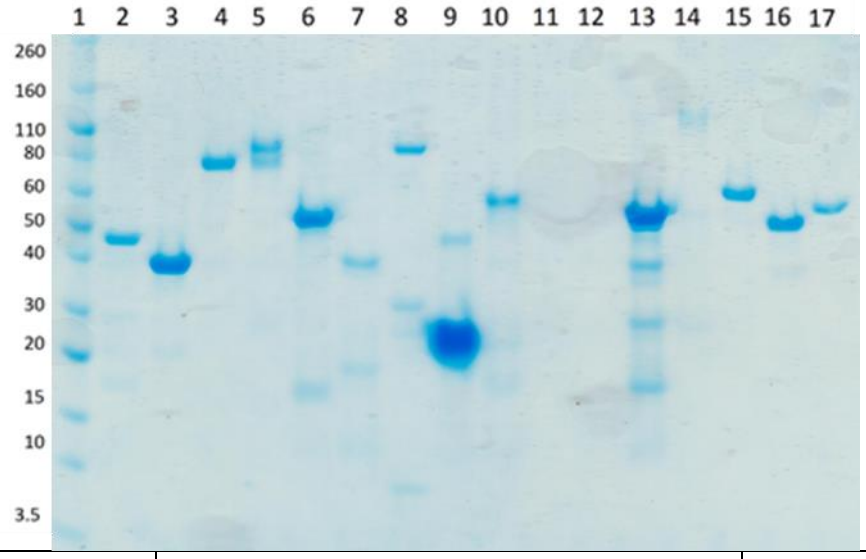
Lane	Protein	MW (kDa)
1	Novex Sharp unstained protein standard	
2	LAGE V1_N-BirA S*	35.9
3	LAGE V1_N-BirA I*	35.9
4	IGF1_N-BirA S*	25.4
5	IGF1_N-BirA I*	25.4
6	IL8_C-BirA S	24.1
7	IL8_C-BirA I	24.1
8	BCL2_C-BirA S	41.1
9	BCL2_C-BirA I	41.1
10	SAA1_N-BirA S*	26.4
11	SAA1_N-BirA I*	26.4
12	AT2R1_N-BirA S*	55.9
13	AT2R1_N-BirA I*	55.9
14	LAGE A1_N-BirA S*	33.1
15	LAGE A1_N-BirA I*	33.1
16	TNF ECD_N-BirA S*	34.4
17	TNF ECD_N-BirA I*	34.4

Figure 17 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project.



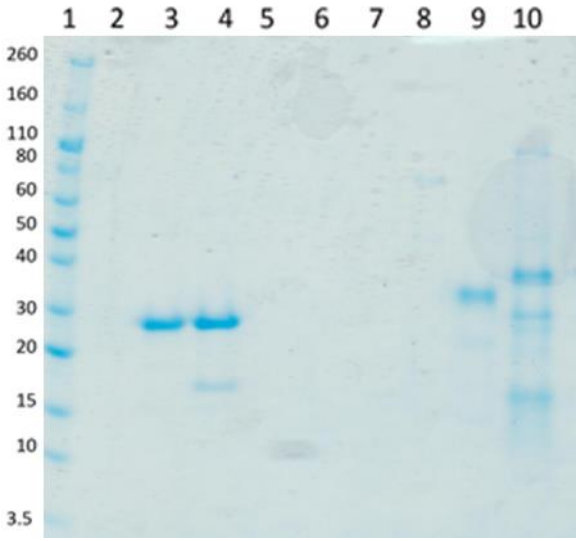
Lane	Protein	MW (kDa)
1	Novex Sharp unstained protein standard	
2	IL8_N-BirA S	24.0
3	IL8_N-BirA I	24.0
4	CDKN1B_C-BirA S	39.9
5	CDKN1B_C-BirA I	39.9
6	TF_N-BirA S	91.9
7	TF_N-BirA I	91.9
8	HOXB7_N-BirA S*	38.8
9	HOXB7_N-BirA I*	38.8
10	IGFBP1_N-BirA S*	40.1
11	IGFBP1_N-BirA I*	40.1
12	APOA1_C-BirA S	43.1
13	APOA1_C-BirA I	43.1
14	HSPA2_N-BirA S	84.9
15	HSPA2_N-BirA I	84.9
16	PRL_C-BirA S	37.9
17	PRL_C-BirA I	37.9

Figure 18 - SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project.



Lane	Protein	MW (kDa)
1	Novex Sharp unstained protein standard	
2	YWHAZ N-BirA I	42.6
3	HNRNPA2 C-BirA I	35.9
4	HNRNPL C-BirA I	76.9
5	MDM2 C-BirA I	70.88
6	AKR1B10 N-BirA I	50.8
7	CRP N-BirA I	37.9
8	HSPA4 N-BirA I	87.4
9	EGF N-BirA I	21.1
10	WT1-2 N-BirA I*	62.1
11	WT1-2N C-BirA I*	40.9
12	WT1-2N C-BirA I*	40.9
13	GIPC1 N-BirA I*	50.9
14	PEGF C-BirA I*	127.2
15	CK18-BirA I	63.6
16	AnnexinII-BirA I	53.3
17	CK19-BirA I	59.3

Figure 19- SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification



Lane	Protein	MW (kDa)
1	Novex Sharp unstained protein standard	
2	CD44 C-BirA I	91.5
3	Annexin I-I C-BirA I*	23.6
4	Annexin I-2 C-BirA I*	23.3
5	CD133-C C-BirA I*	62.7
6	CD133-ECD C-BirA I*	99.4
7	CD133-N C-BirA I*	69.8
8	EGFR1 ECD C-BirA I*	81.8
9	HE-4 M C-BirA I*	25.1
10	XRCC5 N-BirA I	97.5

Figure 20- SDS-PAGE analysis of HTP purified proteins. S=Soluble purification, I= Insoluble purification. *referrers to proteins produced for a separate Oncimmune project.

Table 13 – Protein concentration of each batch of protein chosen for discovery assay

Protein	MW (kDa)	Concentration (µg/ml)
AIF1_N-BirA	32	662
AKR1B10_N-BirA	67	338
AnnexinII_BirA	53	221
APOA1_C-BirA	43.1	294
BCL2_C-BirA	41.1	93
CD44_C-BirA	92	24
CDKN1B_C-BirA	39.9	100
CHGA_N-BirA	63.8	187
CK18_BirA	64	133
CK19_BirA	59	73
CPS1_C-BirA	89.8	198
CPS1_N-BirA	104.9	316
CPS1-C_N-BirA	70	349
DDX3X_C-BirA	88.1	170
DDX3X_N-BirA	88	232
EEF2_N-BirA	110.1	303
EEF2-C_N-BirA	58.2	452
EEF2-N_N-BirA	66.8	316
F2_N-BirA	80.2	252
FUCA1_C-BirA	68.6	0
GLUL_N-BirA	56.8	452
GOLM1_N-BirA	60.2	152
HNRNPA2_C-BirA	36	343
HNRNPL_C-BirA	77	185
HSPA2_N-BirA	84.9	519
HSPA4_N-BirA	87	139
HSPD1_N-BirA	72.9	245
IL6_C-BirA	35.7	709
IL8_C-BirA	24.1	265
IL8_N-BirA	24	916
KRT23_C-BirA	63.1	154
LMNB1_C-BirA	81.3	208
MDM2_C-BirA	71	173
MMP9_C-BirA	81.6	71
NPM1_N-BirA	47.4	305
PEBP1_C-BirA	35.9	225
PRL_C-BirA	37.9	611
RGN_N-BirA	48	341
SALL4B_C-BirA	80.6	70
SALL4-N_C-BirA	55.9	168
SPP1_N-BirA	50.3	187
TF_N-BirA	91.9	314
TGFB1_N-BirA	27.7	252
XRCC5_N-BirA	97	245
YWHAZ_N-BirA	43	152

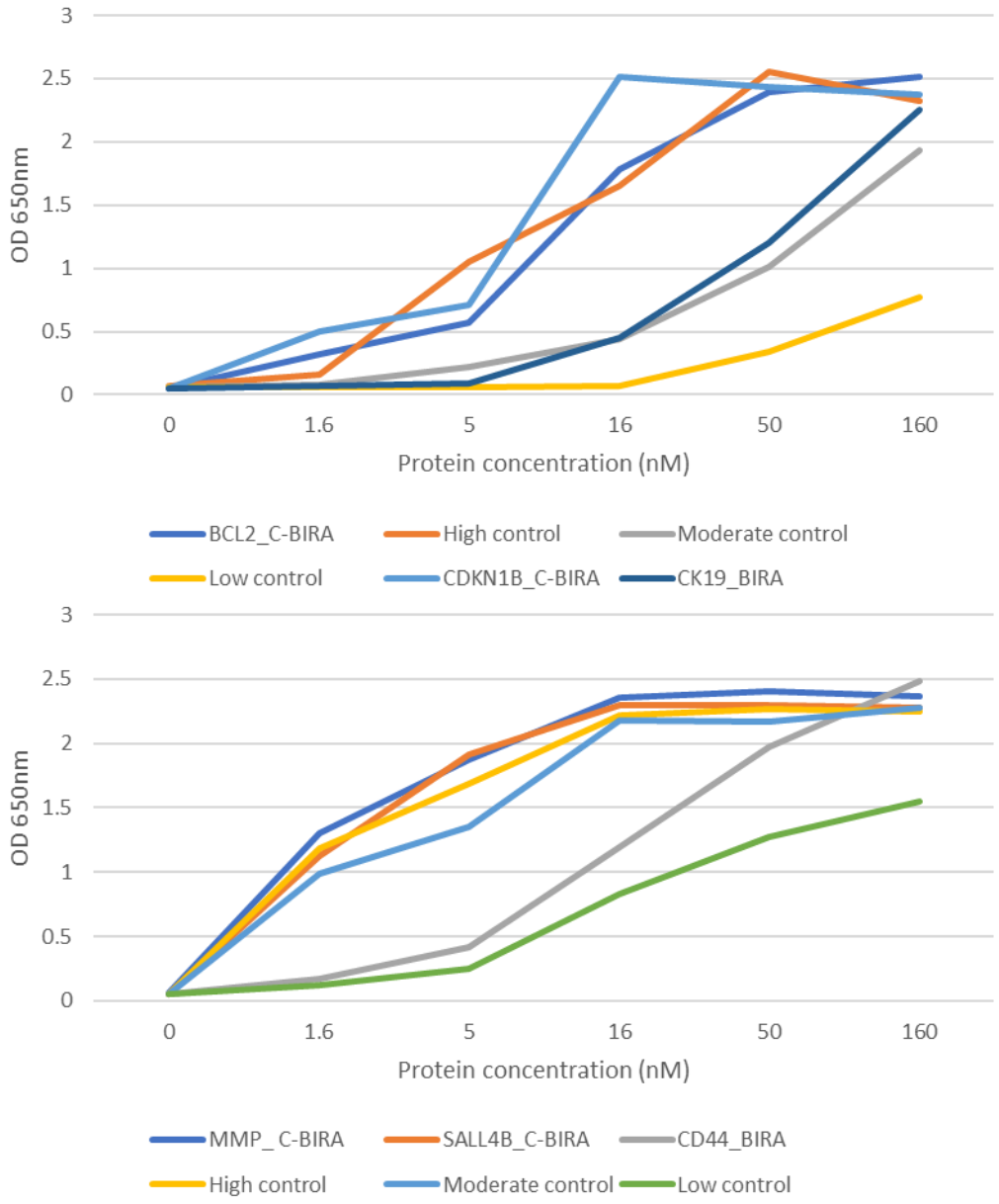


Figure 21- Anti-HIS concentration ELISA for the five batches which showed concentration of 50-100µg/ml by Bradford assay

3.3.3.3 *Full scale expression*

FUCA1_C-BirA and CD44_C-BirA did not produce the required amount of protein from HTP expression so were expressed on a larger scale (section 5.2.1). After 20 hours of expression at 25˚C, CD44_C-BirA cells were harvested by centrifugation, with the resulting pellet weighing 15g. FUCA1_C-BirA was first expressed at 25˚C with a resulting pellet weighing 17g. The 17g pellet yielded very little protein, so expression was repeated at 37˚C. The 37˚C FUCA1_C-BirA pellet weighed 13g.

3.3.3.4 *Full scale purification*

Purification of CD44_C-BirA was performed from the full lysis of the cell pellet (section 5.3.2), from which 1.86mg of protein was obtained at a concentration of 116µg/ml. This was deemed as an acceptable level of protein for the discovery assay. FUCA1_C-BirA was purified from the IB lysate, from which 1.38mg of protein was obtained at a concentration of 284µg/ml and again was deemed acceptable for the discovery assays.

3.4 Discussion

An in-house classification system was designed to assess the evidence of TAAs identified in other studies, in an attempt to select the best of these for analysis in this thesis. 63 TAAs were found in the literature review and testing all of these was beyond the feasibility of this study as another aim was to discover new TAAs. There currently exists no verified system to assess the reliability and performance of TAAs detailed in the literature and therefore this system was created. Firstly, the Youden's J score was considered to assess the ability of the maker to distinguish HCC from controls. This statistic gives a basic measure of the accuracy of a diagnostic test and is suitable for use with the majority of papers which report sensitivity and specificity of a marker at a preselected cut-off. 1 point was assigned for a J score of 0.05-0.1 and 2 points for a score >0.1, as this reflects what has been seen in other cancers with many markers having a score of 0-0.1 and better markers exceeding 0.1[240–242]. The other 3 scores; (I) evidence of early stage sensitivity, (II) the use of appropriate high-risk liver disease controls and (III) the size of the cohort in relation to the number of markers screened, were used to determine the appropriateness and reliability of the study. TAAs with evidence of early stage sensitivity were of interest for this study as this should help to achieve one of the aims of determining a test with an early stage sensitivity greater than AFP. The use of appropriate controls is also considered in this model as this improves the reliability of the data that the AAbs are present as a result of the HCC and not of the underlying liver disease. Assessment of the size of the cohort in relation to the number of markers screened was implemented as a measure of the statistical

appropriateness/power of the study, as increasing the potential markers analysed in a study, without increasing the cohort size, increases the likelihood of markers appearing by chance rather than being genuine results. This issue is often corrected for in microarray based gene expression studies[243], but is rarely mentioned in smaller studies such as the ones analysed in this section.

The primary downside of this classification system is that it is not a validated system and there is no evidence that the scores assigned will correctly weight the total scores to select the right antigens. It may be that an antigen with a truly high Youden's score was excluded from this work due to low scores from the studies reliability and may have been better than any of the antigens selected from more reliable studies. Having this as the only factor to follow-up on studies could ignore new promising leads identified in studies with limited resources. Unfortunately, as all the potential markers were not run in this study, it is impossible to draw a conclusion from the data available. To draw this conclusion, all identified TAAs would need to be produced and tested to determine if the antigens with higher scores made up a higher proportion of the antigens with better differentiation than those with lower scores. Unfortunately, this was not possible within the time-scale and resource restrictions of this thesis, however a future study to fully design and validate a model may be of use to other researchers and future work in other cancers.

With the exception of HGF, cloning proved to be successful. Touchdown PCR was useful in two of the constructs, MMP9-CLIC and KRT23-CLIC. Touchdown PCR helps to increase specificity as only the most specific primer-template annealing will occur at the high temperatures. It also has the advantage of reducing secondary structure formation which may be an issue in high GC content templates.

Gibson assembly was chosen as the preferred cloning method. The homologous regions added to each primer facilitate cloning without the need to incorporate specific restriction enzyme sequences or utilise ligation independent cloning. The entire process can also take place within one reaction, reducing the chance of errors. Gibson assembly creates a knick-free

unmethylated product, allowing the methylated template vector DNA to be digested via DPN1 therefore reducing the number of vector only transformed colonies. Each gene was cloned successfully in at least one construct or variant, with only three mutations occurring, for one of which, CDKN1B-NLIC, there was also a successful CLIC clone. The other two mutations occurred as a triplet deletion in a linker sequence between the antigen insert and the BirA tag. This should have a very minimal effect, if any, so were accepted. A triplet deletion points towards and issue with the primer rather than natural mutations occurring due to the error rate of the polymerase.

HGF cloning proved unsuccessful. The start of HGF is very GC poor, leading to very low annealing temperatures and increased likelihood of secondary structure formation. Secondary structure prediction tools identified the start of HGF to have a very high chance of secondary structure formation, especially in the primer region. After modifying the PCR protocol to reduce secondary structure formation, a product of expected size was produced. Analysis of the post cloning reaction indicated successful cloning and therefore, the problem lies elsewhere. The most likely explanation in this case is the possibility of the DNA/leaky expression of the protein being toxic to the bacteria. This may be solved using different vectors or bacteria strains[244], however time constraints and low priority of the antigen from the literature review led to neglecting this antigen. Protein expression and purification was largely successful.

Chapter 4. Discovery assay

4.1 Introduction

Despite development of new techniques for high-throughput (HTP) AAb discovery, ELISA remains the gold standard. Techniques such as SEREX or Phage-Display have a higher throughput however they often show inconsistent results upon transfer to ELISA for use as a screening test. Using liquid handling robots, it is possible to increase the throughput of a traditional ELISA to test 48 markers per plate, whilst still being able to test a large sample set. In future experiments, this number will be reduced to only the lead antigens run in a dilution series by ELISA and as this will be using the same method, results should be consistent.

4.2 Methods

4.2.1 Serum samples

Serum samples from three groups were used in this assay:

1. 99 HCC samples purchased from Indivumed GmbH
2. 99 samples from patients with non-cancerous chronic liver disease (NCCLD) collected by the Queens Medical Centre biomedical research unit
3. 99 samples from patients with no recorded liver disease or history of cancer collected by Oncimmune ltd.

Demographics of the samples are found in Table 144-Table 166. Underlying liver disease status of the HCC samples was not available.

Table 14 - Age and gender data for discovery assay samples

Demographic factor	HCC patients	NCCLD patients	Individuals with no evidence of malignancy
Number	99	99	99
Mean age	62.3	58.1	62.2
Age range	30-91	30-89	30-87
% Male	69	69	69

Table 15 - HCC stage of discovery assay samples. N/A = stage data unavailable

TNM stage	Number
1	35
2	21
3	21
4	2
N/A	20

Table 16 - Aetiology of the NCCLD controls

NCCLD controls	Number
Hepatitis C viral infection	67
Hepatitis B viral infection	22
Alcoholic Liver Disease	6
Autoimmune Hepatitis	2
Haemochromatosis	1
Primary Biliary Cirrhosis	1

4.2.2 **2µM stock plate preparation**

Antigens were diluted in Dialysis buffer 2 +50% glycerol to a 2µM solution. After diluting each antigen (2µM) they were placed in individual wells of a 48 well stock plate. This stock plate was then used to prepare the 160nm and 50nm stock plates for plate coating (Figure 22).

	1	2	3	4	5	6
A	BLANK	VOL	SSX2_C-BIRA	AIF1_N-BIRA	CHGA_N-BIRA	BCL2_C-BIRA
B	CPS1_N-BIRA	CPS1_C-BIRA	DDX3X_C-BIRA	DDX3X_N-BIRA	APOA1_C-BIRA	PRL_C-BIRA
C	EEF2_N-BIRA	GLUL_N-BIRA	GOLM1_N-BIRA	HSPD1_N-BIRA	KRT23_C-BIRA	LMNB1_C-BIRA
D	PEBP1_C-BIRA	RGN_N-BIRA	F2_N-BIRA	SPP1_N-BIRA	TGFB1_N-BIRA	YWHAZ_N-BIRA
E	HNRNPA2_C-BIRA	HNRNPL_C-BIRA	MDM2_C-BIRA	AKR1B10_N-BIRA	CK18_BIRA	ANXA2_BIRA
F	NPM1_C-BIRA	IL6_C-BIRA	XRCC5_N-BIRA	HSPA4_N-BIRA	SALL4B_C-BIRA	FUCA1_C-BIRA
G	CD44_C-BIRA	CK19_BIRA	MMP9_C-BIRA	VIM_BIRA	EPCAM_ECD-BIRA	HSPA2_N-BIRA
H	CDKN1B_C-BIRA	TF_N-BIRA	IL8_C-BIRA	BLANK	VOL	NY_ESO-1-BirA

Figure 22 - Layout of the 2µM, 160nM and 50nM stock plates

4.2.3 **Plate Coating**

The 2µM stock plate was used to further create 160nM and 50nM stock plates by further diluting antigens in assay coating buffer. Assay plates were created by adding 50µl of each antigen dilution, using a liquid handling robot, to each well of a 96 well plate. The left-hand side of the plate was coated from the 50nm stock plate and right-hand side coated from the 160nm stock plate. Coated 96 well plates were covered and incubated overnight at room temperature for use in the assay the following day. See Figure 23 for the coated plate layout.

	50nM						160nM					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	BLANK	VOL	SSX2_C-BIRA	AIF1_N-BIRA	CHGA_N-BIRA	BCL2_C-BIRA	BLANK	VOL	SSX2_C-BIRA	AIF1_N-BIRA	CHGA_N-BIRA	BCL2_C-BIRA
B	CPS1_N-BIRA	CPS1_C-BIRA	DDX3X_C-BIRA	DDX3X_N-BIRA	APOA1_C-BIRA	PRL_C-BIRA	CPS1_N-BIRA	CPS1_C-BIRA	DDX3X_C-BIRA	DDX3X_N-BIRA	APOA1_C-BIRA	PRL_C-BIRA
C	EEF2_N-BIRA	GLUL_N-BIRA	GOLM1_N-BIRA	HSPD1_N-BIRA	KRT23_C-BIRA	LMNB1_C-BIRA	EEF2_N-BIRA	GLUL_N-BIRA	GOLM1_N-BIRA	HSPD1_N-BIRA	KRT23_C-BIRA	LMNB1_C-BIRA
D	PEBP1_C-BIRA	RGN_N-BIRA	F2_N-BIRA	SPP1_N-BIRA	TGFB1_N-BIRA	YWHAZ_N-BIRA	PEBP1_C-BIRA	RGN_N-BIRA	F2_N-BIRA	SPP1_N-BIRA	TGFB1_N-BIRA	YWHAZ_N-BIRA
E	HNRNPA2_C-BIRA	HNRNPL_C-BIRA	MDM2_C-BIRA	AKR1B10_N-BIRA	CK18_BIRA	ANXA2_BIRA	HNRNPA2_C-BIRA	HNRNPL_C-BIRA	MDM2_C-BIRA	AKR1B10_N-BIRA	CK18_BIRA	ANXA2_BIRA
F	NPM1_C-BIRA	IL6_C-BIRA	XRCC5_N-BIRA	HSPA4_N-BIRA	SALL4B_C-BIRA	FUCA1_C-BIRA	NPM1_C-BIRA	IL6_C-BIRA	XRCC5_N-BIRA	HSPA4_N-BIRA	SALL4B_C-BIRA	FUCA1_C-BIRA
G	CD44_C-BIRA	CK19_BIRA	MMP9_C-BIRA	VIM_BIRA	EPCAM_ECD-BIRA	HSPA2_N-BIRA	CD44_C-BIRA	CK19_BIRA	MMP9_C-BIRA	VIM_BIRA	EPCAM_ECD-BIRA	HSPA2_N-BIRA
H	CDKN1B_C-BIRA	TF_N-BIRA	IL8_C-BIRA	BLANK	VOL	NY_ESO-1-BirA	CDKN1B_C-BIRA	TF_N-BIRA	IL8_C-BIRA	BLANK	VOL	NY_ESO-1-BirA

Figure 23 - Layout of the 96 well plates. One 96-well plate was used per sample

4.2.4 High Throughput Assay protocol

All serum samples were diluted 1 in 110 assay blocking buffer, before vortexing to ensure homogeneity. Samples were kept at 4˚C until required.

Antigen coated plates (see plate layout in section 4.2.3) were washed three times with assay wash buffer (3x 200µl per well) and subsequently blocked with assay blocking buffer for one hour with PSGBB (200µl per well). After blocking, plates were aspirated and diluted serum (50µl per well) was dispensed into each well of a coated 96 well plate, using one plate per patient serum sample. Two plates were prepared as assay controls also; a blank plate (no serum) was incubated with assay blocking buffer and a plate incubated with mouse anti-HIS IgG (Abcam, Ab15149, 50µl per well). The plates were incubated, wrapped in saran wrap, with shaking at 400rpm for 90 minutes at room temperature (RT). Plates were washed with assay wash buffer (3x 200µl per well) prior to incubation with rabbit anti-human IgG-HRP (Dako, 50µl per well), with shaking (400rpm, 1 hour RT). Plates were washed again (3x assay wash buffer, 200µl per well) before addition of 3, 3´, 5, 5´-tetramethylbenzidine (TMB, 50µl per well) and incubated without shaking (15 minutes, RT) and the OD of each well was measured (650nm).

4.2.5 Verification of plate layout

4.2.5.1 *Coomassie stained SDS-PAGE*

To verify the correct plate layout of antigens 2µM stock plate were analysed by Coomassie stained SDS-PAGE gels and anti-HIS tag western blots. Antigen solutions (2µM) were diluted (1:1, dH_20) prior to addition of DTT and LDS sample buffer. Coomassie staining was performed as described in section 3.2.5.4**Error! Reference source not found.**.

4.2.5.2 *Western blot*

For western blots, proteins were prepared as in section 4.2.5.1 and subsequently run on 4-20% gradient Precast RunBlue SDS-PAGE gels (Expedeon) for 45 minutes at 170v. Proteins were transferred to nitrocellulose

membranes using RunBlue TG blot buffer at 10v for 16 hours overnight and subsequently blocked in 5% milk in TBST (25ml, 1 hour). Blots were probed with anti-HIS IgG (Abcam ab15149) diluted in 5% milk (10ml, 1 hour), washed in TBST (3x 5mls) and probed with a HRP conjugated rabbit anti-mouse IgG (Dako, 1 hour). ECL detection reagent and Hyperfilm ECL (Amersham) were used for western blot development, with a 5-minute exposure time before development and fixation in developer/fixation solutions (Kodak).

4.2.6 **AFP measurement**

Circulating alpha-fetoprotein (AFP) was measured in the serum samples using a commercially available ELISA (Aviva Systems Biology) following the manufacturer's instructions.

4.2.7 **Data analysis**

AAb response ODs for each antigen at 160nM and 50nM were normalised against the signal from the corresponding vector only lysate (VOL) well by subtracting the VOL signal from the AAb signal. VOL is the expression product of the BirA solubility tag including the 6xHIS tag. Cut-offs for individual AAbs were set for 160nm and 50nm dilutions as the value which gave the highest Youden's J statistic at a specificity of >=95%. Panels were formed using Net reclassification improvement[245]. This analysis works on the principle of adding in AAbs one at a time determined by which AAb would add the most to the overall accuracy of the panel. After addition of an AAb the NRI score is recalculated before adding the next AAb until no further improvement in accuracy can be obtained. AFP was given a cut-off of 200ng/ml, this is a commonly used cut-off for AFP and avoids the low specificity which is often seen at 20ng/ml, especially in HCV+ patients[246]. Accuracy of tests if calculated as:

$$\frac{True\ positives + True\ negatives}{True\ positives + True\ negatives + False\ positives + False\ negatives}$$

Significance analysis for difference sensitivity of the panel at individual stages of tumour was assessed using McNemar's test. The Stata (Stata corp.) package roccomp was used to assess accuracy improvements between tests.

4.3 Results

4.3.1 Stock plate confirmation

Stock plates were created successfully except for well G6, containing HSPA2_N-BirA (see Figure 22), which was contaminated with CDKN1B_C-BirA. This led to well H1 CDKN1B_C-BirA being much weaker than intended as there was less protein remaining than needed to form a 2µM solution. HSPA2_N-BirA was loaded into well H4 instead of an assay buffer only well. Coomassie stained gels and western blots are shown in Figure 24 & 25. As these proteins were prepared from an equimolar stock plate, you would expect bands in the Coomassie stained gel to show the same intensity. These vary however as the protein concentration was based on total protein, so higher levels of contaminating proteins will lead to a lower intensity of band at the correct size. Some proteins were not purified at a high enough concentration to create a 2µM stock dilution, however if the intended 160nM dilution would create a dilution of at least 50nM they were deemed acceptable. In this situation, the band would also be weaker.

The correct size protein can be seen in each lane from either the Coomassie gel or the western blot for all but six proteins. For FUCA1_C-BirA, a band shows on both the Coomassie and the anti-HIS western blot approximately 10kDa below its predicted value. This is consistent with all previous gels produced of this protein, this is likely the correct protein as it is not uncommon for proteins to run +/- 10kDa different to predicted[247]. RGN_N-BirA, XRCC5_N-BirA and CD44_C-BirA could not be confirmed by either Coomassie staining or western blot which is surprising as both RGN_N-BirA and XRCC5_N-BirA were at full 2µM concentrations. The N-BirA constructs do not react well with the anti-HIS antibody, most likely due to slightly different folding of the tag. As western blot is more sensitive than Coomassie staining, low levels of these proteins not showing on Coomassie will not be detected on anti-HIS either due to the N-term His tag issue. XRCC5_N-BirA showed signs of degradation when produced and is therefore possible it has degraded further, this could explain why there is no longer a visible band. RGN_N-BirA is surprising as the original production

QC showed a strong band with little sign of degradation. A band of the predicted size of CDKN1B_C-BirA can be seen in the contaminated well, this helps to confirm that the contaminated well was correctly identified.

Lane	Gel A	Mol. weight (kDa)	Lane	Gel B	Mol weight (kDa)	Lane	Gel C	Mol weight (kDa)
1	Novex Sharp ladder		1	Novex Sharp ladder		1	Novex Sharp ladder	
2	Blank	0	2	KRT23_C-BirA	63	2	XRCC5_N-BirA	97
3	VOL	16	3	LMNB1_C-BirA	81	3	HSPA4_N-BirA	87
4	SSX2_C-BirA	37	4	PEBP1_C-BirA	36	4	SALL4B_C-BirA	81
5	AIF1_N-BirA	32	5	RGN_N-BirA	48	5	FUCA1_C-BirA	68.6
6	CHGA_N-BirA	64	6	F2_N-BirA	80	6	CD44_C-BirA	91.5
7	BCL2_C-BirA	41	7	SPP1_N-BirA	50	7	CK19_BirA	59.3
8	CPS1_N-BirA	105	8	TGFB1_N-BirA	28	8	MMP9_C-BirA	82
9	CPS1_C-BirA	90	9	YWHAZ_N-BirA	43	9	VIM_BirA	69.2
10	DDX3X_C-BirA	88	10	HNRNPA2_C-BirA	36	10	EPCAM_ BirA	45
11	DDX3X_N-BirA	88	11	HNRNPL_C-BirA	77	11	contaminated well	0
12	APOA1_C-BirA	43	12	MDM2_C-BirA	71	12	CDKN1B_C-BirA	40
13	PRL_C-BirA	38	13	AKR1B10_N-BirA	67	13	TF_N-BirA	92
14	EEF2_N-BirA	110	14	CK18_BirA	64	14	IL-8_C-BirA	24
15	GLUL_N-BirA	57	15	AnnexinII_BirA	53	15	HSPA2_N-BirA	85
16	GOLM1_N-BirA	60	16	NPM1_C-BirA	47	16	VOL	16
17	HSPD1_N-BirA	73	17	IL6_C-BirA	36	17	NY-ESO-1_BirA	34

Figure 24 – Legend for figure 25, showing the confirmation Coomassie stained gels and anti-HIS westerns of the stock plate layout.

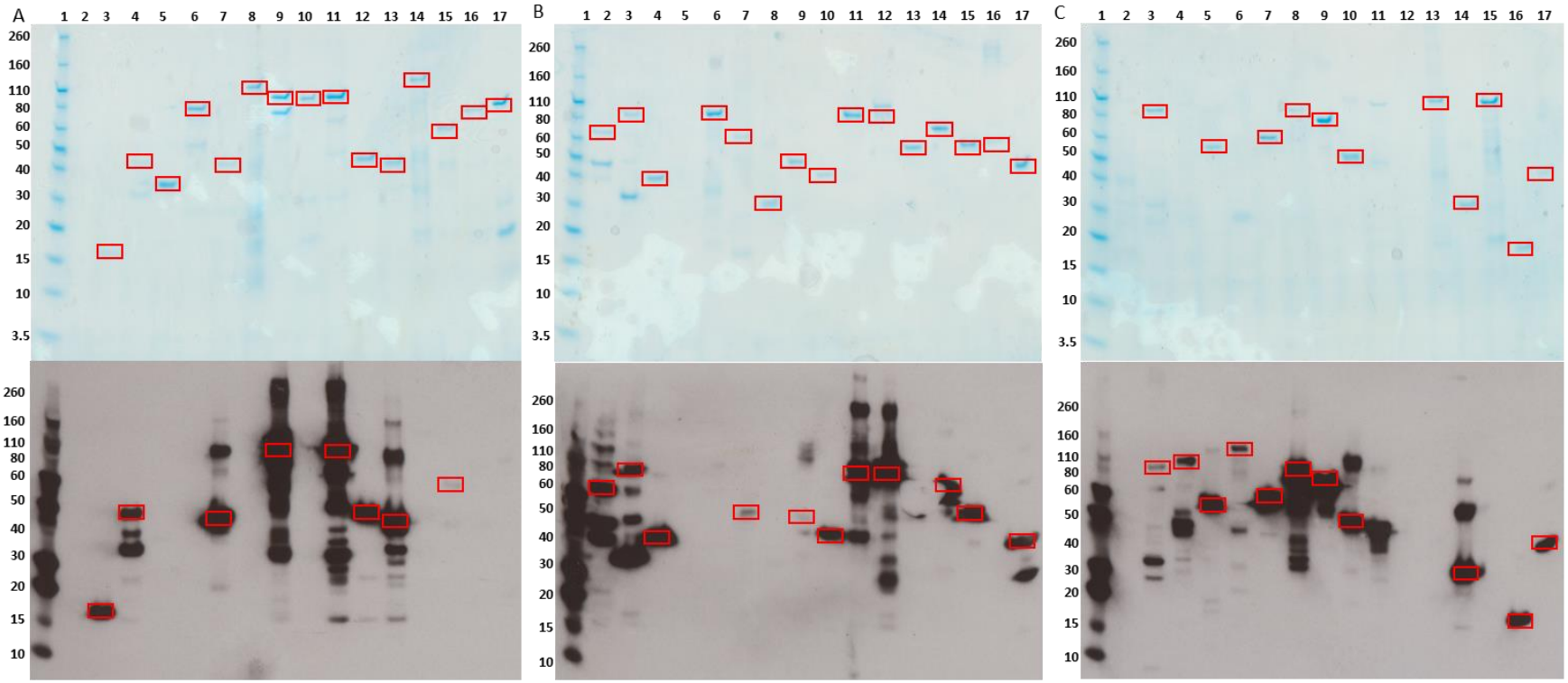


Figure 25 –Coomassie stained gels (top) and anti-HIS westerns (bottom) of the stock plate layout. Red boxes highlight the band at the expected size

4.3.2 **High Throughput Assay results**

4.3.2.1 ***Individual AAb performance***

After VOL correction, cut-offs were chosen for each individual AAb at the two assayed concentrations. The number of positive samples in each group was then calculated from these cut-offs (Table 177). Example dot plots of some of the AAbs by group are shown in Figure 26. NY-ESO-1_BirA was the best performing AAb from this screen detecting seven cancers with zero false positives in the NCCLD group (Youden's score of 0.07) and only one false positive in the healthy control group. NY-ESO-1 had previously shown ability to differentiate between HCC and NCCLD samples[170] so was therefore included as a positive control. HSPA4_N-BirA had the best Youden's value of the new AAbs (0.04), detecting four HCCs with zero false positives in either group at a concentration of 50nM. No single AAb would be suitable as a screening tool for detecting HCC in either a high-risk liver disease population or healthy population.

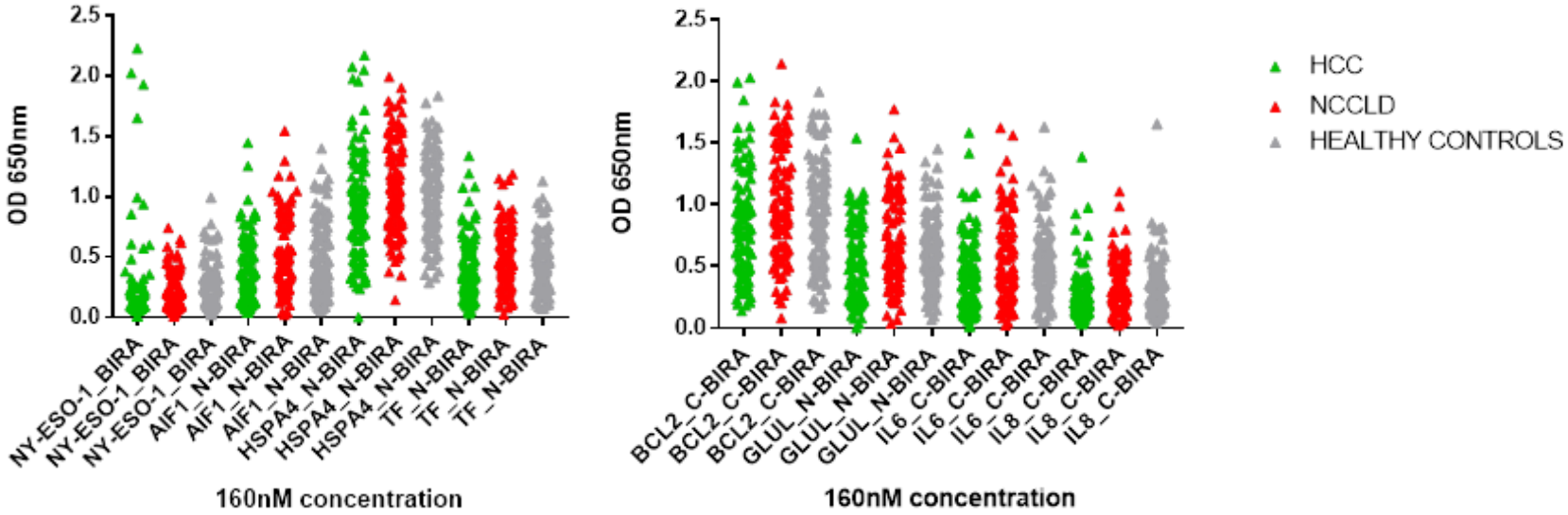


Figure 26 – Dot plot results of a selection of AAbs run in this study split by sample type

Table 17 – Number of positive samples in each sample group for each AAb

AAb	HCC Positives	NCCLD Positives	Healthy control Positives
AIF1_N-BirA 50nM	2	0	0
AIF1_N-BirA 160nM	2	2	1
AKR1B10_N-BirA 50nM	1	0	0
AKR1B10_N-BirA 160nM	1	1	0
AnnexinII_BirA 50nM	1	1	0
AnnexinII_BirA 160nM	1	1	0
APOA1_C-BirA 50nM	1	0	0
APOA1_C-BirA 160nM	1	1	0
BCL2_C-BirA 50nM	1	1	0
BCL2_C-BirA 160nM	3	1	1
CD44_C-BirA 50nM	1	0	0
CD44_C-BirA 160nM	1	0	0
CDKN1B_C-BirA 50nM	2	0	1
CDKN1B_C-BirA 160nM	7	4	6
CHGA_N-BirA 50nM	1	1	1
CHGA_N-BirA 160nM	1	1	1
CK18_BirA 50nM	1	0	0
CK18_BirA 160nM	1	1	0
CK19_BirA 50nM	1	4	1
CK19_BirA 160nM	1	3	1
CPS1_C-BirA 50nM	1	1	0
CPS1_C-BirA 160nM	2	1	0
CPS1_N-BirA 50nM	1	1	0
CPS1_N-BirA 160nM	1	1	0
DDX3X_C-BirA 50nM	1	1	0
DDX3X_C-BirA 160nM	1	1	0
DDX3X_N-BirA 50nM	4	4	3
DDX3X_N-BirA 160nM	3	1	2
EEF2_N-BirA 50nM	1	1	0
EEF2_N-BirA 160nM	1	1	0
EPCAM_BirA 50nM	5	1	6
EPCAM_BirA 160nM	2	0	1
F2_N-BirA 50nM	1	1	0
F2_N-BirA 160nM	1	2	3
FUCA1_C-BirA 50nM	1	0	0
FUCA1_C-BirA 160nM	1	3	0
GLUL_N-BirA 50nM	1	0	0
GLUL_N-BirA 160nM	1	2	0
GOLM1_N-BirA 50nM	1	1	1
GOLM1_N-BirA 160nM	2	2	2
HNRNPA2_C-BirA 50nM	1	0	0
HNRNPA2_C-BirA 160nM	1	0	0
HNRNPL_C-BirA 50nM	2	3	0
HNRNPL_C-BirA 160nM	3	2	0
HSPA2_N-BirA 50nM	4	5	3
HSPA2_N-BirA 160nM	1	0	0
HSPA4_N-BirA 50nM	4	0	0

HSPA4_N-BirA 160nM	5	1	0
HSPD1_N-BirA 50nM	2	2	0
HSPD1_N-BirA 160nM	2	2	0
IL-6_C-BirA 50nM	2	2	2
IL-6_C-BirA 160nM	2	2	1
IL-8_C-BirA 50nM	5	4	4
IL-8_C-BirA 160nM	4	2	5
KRT23_C-BirA 50nM	1	1	0
KRT23_C-BirA 160nM	1	1	2
LMNB1_C-BirA 50nM	1	1	0
LMNB1_C-BirA 160nM	2	2	2
MDM2_C-BirA 50nM	2	1	0
MDM2_C-BirA 160nM	2	1	1
MMP9_C-BirA 50nM	1	1	0
MMP9_C-BirA 160nM	3	1	1
NPM1_C-BirA 50nM	1	0	0
NPM1_C-BirA 160nM	2	1	0
NY-ESO-1_BirA 50nM	8	1	3
NY-ESO-1_BirA 160nM	7	0	1
PEBP1_C-BirA 50nM	1	0	0
PEBP1_C-BirA 160nM	1	0	0
PRL_C-BirA 50nM	1	0	0
PRL_C-BirA 160nM	2	1	1
RGN_N-BirA 50nM	1	0	0
RGN_N-BirA 160nM	1	0	0
SALL4B_C-BirA 50nM	1	0	0
SALL4B_C-BirA 160nM	2	2	0
SPP1_N-BirA 50nM	1	1	0
SPP1_N-BirA 160nM	1	1	1
SSX2_C-BirA 50nM	3	1	1
SSX2_C-BirA 160nM	3	2	3
TF_N-BirA 50nM	2	1	1
TF_N-BirA 160nM	2	0	0
TGFB1_N-BirA 50nM	1	0	0
TGFB1_N-BirA 160nM	2	2	1
VIM_BirA 50nM	3	1	4
VIM_BirA 160nM	7	3	6
XRCC5_N-BirA 50nM	1	1	0
XRCC5_N-BirA 160nM	1	2	0
YWHAZ_N-BirA 50nM	1	1	0
YWHAZ_N-BirA 160nM	2	1	2

4.3.2.2 *Panel formation*

Although no individual AAb shows enough promise to form a screening test on its own, multiple examples of using a panel of AAbs as a screening/diagnostic test exist in the literature[170,182,241,248,249]. Panels were built using Net Reclassification Improvement (NRI) analysis. Table 188 shows the panel performance at the addition of each AAb, ultimately resulting in a panel with a sensitivity of 28% at a specificity of 94% in the NCCLD group and 88% in the healthy controls. This improves from the accuracy of NY-ESO-1_BirA alone, for separating HCC from NCCLD patients, from 54% to 61%.

Table 18 – NRI formation of the AAb panel, sensitivity and specificity values shown are cumulative as each marker is added to the panel

Addition order	Model	NRI score	Sensitivity (%)	Specificity NCCLD (%)	Specificity healthy controls (%)
1	NY-ESO-1_BirA 160nM	0.07	7.07	100.00	98.99
2	EPCAM_BirA 50nM	0.04	12.12	98.99	92.93
3	HSPA4_N-BirA 50nM	0.03	15.15	98.99	92.93
4	VIM_BirA 160nM	0.03	21.21	95.96	88.89
5	HNRNPL_C-BIRA 160nM	0.02	24.24	94.95	88.89
6	TF_N-BirA 160nM	0.01	25.25	94.95	88.89
7	AIF1_N-BirA 50nM	0.01	26.26	94.95	88.89
8	MMP9_C-BirA 160nM	0.01	28.28	93.94	87.88

4.3.2.3 *AFP addition*

A commonly used high specificity cut-off of 200ng/ml was used for AFP. AFP was able to distinguish HCC with a sensitivity of 35% and a specificity of 100% in both control groups. This results in an accuracy of 68% in either control group. When combined with the AAb panel, the sensitivity increased to 53% with no reduction in specificity in the NCCLD control group or the healthy control group. This results in improved accuracy against either control group. The accuracy against the NCCLD control group, the real-life situation control group, demonstrates a significant increase from 68% for AFP alone to 74% for the AAb panel plus AFP (p=0.0102).

4.3.2.4 *Performance by stage*

Table 19 shows performance by stage of the AAb panel, AFP, and in combination. The AAb panel has a higher sensitivity than AFP in stage one HCCs, although not significant (p=0.564), whereas AFP has a higher sensitivity in the remaining stages. This is unsurprising given literature evidence which shows that AFP levels are often associated with stage of the tumour. Combination of the two tests shows an increase in sensitivity of every stage and a significant increase in sensitivity for stage one HCCs (p=0.0143).

Table 19 – Performance of the AAb panel, AFP and the combination of both by HCC stage

Stage	Positives	Negatives	Sensitivity (%)
AAb Panel			
1	8	27	23
2	4	17	19
3	8	13	38
4	0	2	0
N/A	8	12	40
AFP			
1	6	29	17
2	8	13	38
3	11	10	52
4	1	1	50
N/A	9	11	45
AAb Panel+AFP			
1	12	23	34
2	10	11	48
3	14	7	67
4	1	1	50
N/A	14	6	70

4.4a

The findings in these experiments, that no AAb is solely useful in distinguishing HCC from high risk controls, is not surprising as shown previously by ourselves and many others[182,194,250,251]. The heterogeneity of HCC could explain why no individual marker performs well enough and supports the use of a panel of

markers instead. This has been seen in multiple markers, including AFP, DCP, AFP-L3 (see section 1.2.3), not just in AAbs.

The primary aim of this project was to add to/improve upon the existing AAb panel developed by Oncimmune, due to this, NY-ESO-1_BirA, although part of the panel developed previously[170], was used in panel formation for experiment as it appears to be the best performing AAb marker so will most likely remain in the final improved panel. Therefore, AAbs that are additive to NY-ESO-1 are likely useful, in this sample set seven AAbs from were additive to NY-ESO-1 alone, generating a panel with 28% sensitivity at a specificity of 94% in the NCCLD group and 88% in the healthy controls. This panel was built to distinguish HCC from NCCLD, however it is surprising to see a large reduction in specificity in the healthy controls, this is potentially caused by the difference in age between the healthy controls and NCCLD control samples (mean age 58.1 years vs. 62.2 years). In future studies, closer age matching should be achieved to address this issue.

AFP performance at a cut-off of 200ng/ml in this sample set was within the ranges reported in the literature[252–255]. Performance of AFP was worse than the AAb panel in stage one HCCs but improved for other stages. It is commonly seen for AFP to have low sensitivity in early stage cancers. Whilst the difference in sensitivity between the two panels was not significant, AAb addition significantly increased the sensitivity of AFP in stage one HCCs. Using AFP at this single cut-off is a potential limitation of this experiment, as there may be a more optimal cut-off to be used in combination with the AAb panel, potentially leading to increased sensitivity. However, using AFP at this high specificity cut-off, as opposed to a high sensitivity cut-off such as 20ng/ml, should help to overcome any bias towards AFP present in the sample set. As this sample set was collected retrospectively and involved commercially purchased samples, it is hard to say whether AFP was used in the clinical workup of the samples. The cancer samples are likely to be skewed towards higher AFP levels, as those patients are more likely to be diagnosed with HCC when AFP is measured as part of the clinical workup. On the alternative side, it is likely that in the NCCLD

samples, AFP values will be lower as slightly raised AFP levels would infer suspicion of HCC and would therefore not be chosen as samples to use as non-cancerous controls.

The eight leads identified in this panel will be produced on a larger scale and retested on the same sample set. This is necessary to check that the positive samples from this study are specific signals rather than binding to *E. coli* contamination in the HTP protein preparations. Scaled-up protein purification using fast protein liquid chromatography (FPLC) produces purer batches of protein; however, it also requires more time and resources. As well as the eight leads from the panel, six other AAbs which showed HCC/NCCLD differentiation will also be retested, this will allow for some redundancy as certain samples are positive for more than one AAb. The 14 AAbs to be scaled up and retested will be; HSPA4, VIM, EPCAM, TF, AIF1, MMP9, HNRNPL, CDKN1B, HSPD1, HNRNPA2, SALL4B, NPM1, YWHAZ and DDX3X.

Chapter 5. Lead antigen reproduction

5.1 Introduction

Whilst the HTP expression and purification protocols achieved sufficient level of antigens in quality and quantity for initial AAb identification, it is not ideal for confirmation of the results and the long-term use in the assay. As every culture is treated the same during purification, optimisation of the process for purity and yield cannot be achieved at this stage. The use of a FPLC system, such as the AKTA Prime plus (GE Healthcare), allows continuous monitoring of the purification process. The major advantage of this system is the ability to detect what is flowing through the affinity column in real time. Using this system, wash buffer can be applied after the initial lysate load until the flow-through trace returns to zero (Figure 27). This significantly reduces the levels of contamination in the final product, as there should be little non-tagged protein remaining. The effect of this can be seen in a screen of NY-ESO-1_BirA levels in ovarian cancer patients (Figure 28). Of the leads identified in section Chapter 4, 12 required remaking as only HTP batches of antigen had been made.

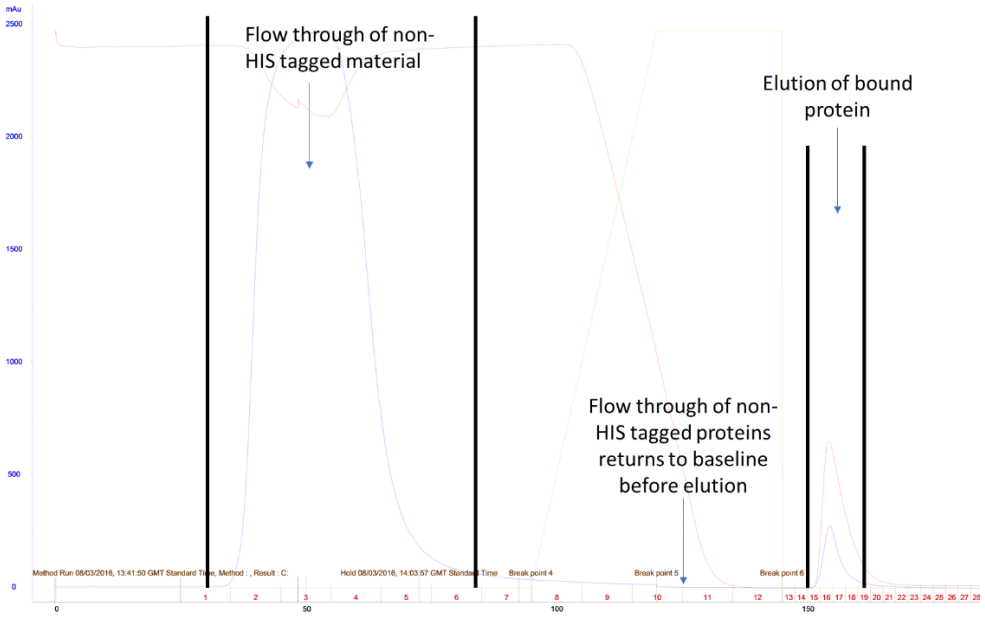


Figure 27 - Example chromatogram from an AKTA prime plus. Blue line shows the UV measurement of the liquid leaving the column and is representative of the protein content of the liquid

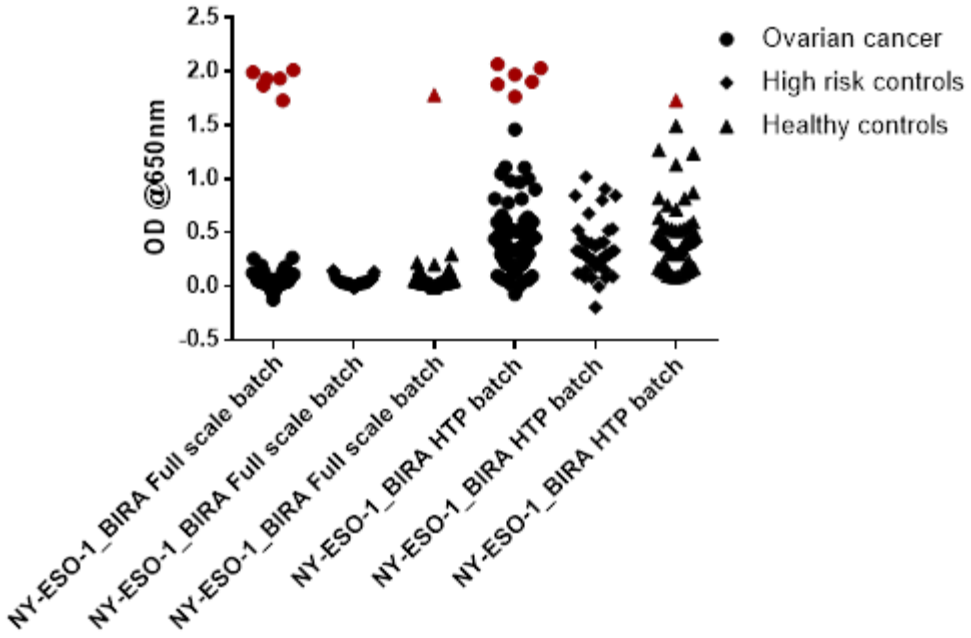


Figure 28 – Dot plot of AAb reactivity of three groups of samples, ovarian cancer, high risk controls and healthy controls against two batches of antigen. One batch (HTP) was produced using the HTP technique described in chapter Chapter 4 and the full-scale batch produced using the technique described in section Chapter 5. Samples highlighted in red were positive with either batch of antigen.

5.2 Methods

5.2.1 Expression

5.2.1.1 *Transformation*

See section 3.2.3.1

5.2.1.2 *Starter cultures*

To create starter cultures, 10ml of 2xYT media plus relevant antibiotics and 1% glucose was inoculated with three colonies from the transformation plate and incubated at 37°C for 6 hours with shaking at 250rpm. After 6 hours growth, the 10ml culture was added to 190ml of the same media in a 2.5L Thompson flask and grown overnight for 16 hours. After 16 hours growth, OD and pH were measured before inoculation of the expression culture.

5.2.1.3 *Expression cultures*

Starter cultures were used to inoculate 500ml of TBOnex-P so that the resulting OD was 0.2. Expression cultures were at incubated 37°C for 2 hours, then 25°C for 18 hours with shaking at 250rpm. After incubation cells were harvested by centrifugation (6000g, 4°C, 30 minutes), the supernatant discarded, and remaining cell pellets frozen at -80°C until purification. A 1ml sample of culture was collected at induction and at harvesting for expression analysis. The 1ml samples were centrifuged (13,000 rpm, 1 minute), the supernatant discarded and remaining pellets stored at -80°C until use.

5.2.1.4 *Expression analysis*

The 1ml expression samples were lysed in two stages. Firstly, soluble lysis was accomplished by resuspending the pellet in soluble lysis buffer and rotated (45rpm, 20 minutes). Tubes were centrifuged (13,000rpm, 10 minutes), the supernatant was collected, identified as the soluble lysate and stored at 4°C. Subsequently 300µl of urea binding buffer was added to the remaining pellet and rotated (45rpm, 20 minutes). After centrifugation (13,000rpm, 10 minutes) the supernatant was collected, identified as the IB lysate and stored at 4°C.

Both the Soluble and IB lysates were analysed by Coomassie staining and Western blot (section 3.2.5.4.1 and 4.2.5.2). For Coomassie gels, samples were added neat to the LDS and SDS mixture, whereas samples were diluted 1 in 10 before addition to the LDS and SDS mixture for Western blot.

5.2.2 Purification

5.2.2.1 *Lysis*

For the soluble lysate, soluble lysis buffer was added at a ratio of 5ml per gram of bacterial pellet before homogenisation via stirring (20 minutes). After centrifugation (11,000g, 4˚C, 20 minutes), the supernatant was transferred to a separate vial and the pellet resuspended in soluble lysis buffer at 2.5ml per gram of the initial bacterial pellet. The homogenisation and centrifugation steps were repeated, and the supernatant added to the lysate from the 5ml per gram. This was used as the soluble lysate for purification. The remaining pellet (IB) was either further lysed or discarded.

If the IB lysate was required for purification, the final pellet from the steps described above, was washed one additional time with a 1 in 10 dilution of soluble lysis buffer and distilled water (dH_2O) and centrifuged (11,000g, 4˚C, 20 minutes). The supernatant was discarded and the remaining IB pellet was resuspended in 2mls per gram of the original pellet weight of IB solubilisation buffer and homogenised by magnetically stirring overnight. After centrifugation (11,000g, 4˚C, 20 minutes), the supernatant was stored as the IB lysate. The remaining pellet was discarded.

If full lysate was required for purification, the cell pellet was resuspended in 5ml per gram of IB lysis buffer and homogenised by magnetically stirring overnight. After centrifugation (11,000g, 4˚C, 20 minutes), the supernatant was removed and stored as the full lysate. The remaining pellet was discarded.

Before purification could take place, the selected lysate was filtered through 0.2µm sterile filter membranes.

5.2.2.2 *Affinity purification*

Separate buffer lines on an Akta Prime Plus system were equilibrated with the following buffers; dH_2O, guanidine binding buffer (GBB), urea binding buffer (UBB), urea elution buffer (UEB) and urea column scrubbing buffer (USCB). All buffers used with the Akta Prime plus during purification were filtered through 0.2µm filters. A 5ml His-trap FF crude column (GE Healthcare) was connected and equilibrated first in dH_2O and then GBB. Sample was applied at a rate of 2.5ml per minute. After sample application, the column was washed with GBB until the UV_{280} trace reached <10mAU. During sample application, the flow-through was collected in 10ml fractions. The column was then buffer exchanged into UBB as a gradient from 0%-100% over 5-column volumes (CV) followed by a 5-CV wash with UBB. Bound proteins were eluted with 5 CV of UEB, with fractions collected every 2.5ml, followed by 5 CV of UCSB. Re-equilibration of the column was performed by running 5 CV of GBB. Subsequent runs were performed by pooling all the flow-through fraction with UV_{280}>100mAU, from the preceding run and repeating the steps above. This was performed until the elution peak reached <50mAU (pellet exhaustion). After each run, elution fractions with a UV_{280} of >50mAU were pooled to form the elution from that run and stored at 4˚C. After pellet exhaustion, all elution pools were combined. A 1ml sample of the final elution pool was saved at 4˚C for quantification and characterisation purposes.

5.2.2.3 *Dialysis*

Protein refolding and Urea/Imidazole removal was performed via dialysis. All steps were performed with gentle agitation and at 4˚C. The combined elution pool was loaded into dialysis tubing with an 8000Da cut-off. This tubing was then placed in the appropriate buffer for dialysis, starting with 5x the volume of the sample of 4M dialysis buffer (4MDB) for 4 hours, followed by 10x volume of 2M Dialysis buffer (2MDB) for 16 hours, 10x volume of 1M dialysis buffer (1MDB) for 8 hours and 10x volume of urea dialysis buffer 2 (UDB2) for 16 hours. A 1ml sample of the final desalted and refolded protein was collected and kept at 4˚C for quantification and characterisation.

5.2.2.4 *Quantification and characterisation*

Protein concentration was assessed by Bradford assay for the elution pool and the desalted samples (section 3.2.5.4.2). Samples from the elution pool and the refolded/desalted sample were assessed for correct molecular weight via anti-HIS tag western blot as described in section 4.2.5.2 with the following modifications:

- SDS-PAGE was performed using 12-4% gradient Novex Bis-Tris Precast Protein gels (Invitrogen) run for 60 minutes at 170v
- Gels were loaded with 1µg of denatured sample per lane

5.2.2.5 *Protein storage*

After quantification and characterisation of the purified proteins, an equal volume of 90% glycerol was added to bring the final glycerol concentration to 50%, followed by DTT and EDTA to a final concentration of 1mM DTT and 0.1mM EDTA and stored at -80˚C. Final storage concentration was recorded as 50% of the desalted sample concentration.

5.3 Results

5.3.1 Expression

Following the discovery assay, 11 leads required scale-up. Every flask produced protein of the correct molecular weight under these conditions. For all bar two proteins, AIF1 N-BirA and MMP9 C-BirA, most of the protein was present in the 20-hour IB sample. For AIF1 N-BirA and MMP9 C-BirA which had a less distinguishable band in the IB sample, full lysis was performed. Final pellet weight and chosen lysis methods are shown in Table 20.

Table 20 – Pellet weight and lysis method chosen for each antigen

Antigen	Lysis	Pellet weight (g)
AIF1_N-BIRA	Full lysis	15
CDKN1B_C-BIRA	IB	19
DDX3X_N-BIRA	IB	14
HNRNPA2_C-BIRA	IB	17
HNRNPL_C-BIRA	IB	16
HSPA4_N-BIRA	IB	16
HSPD1_N-BIRA	IB	13
MMP9_C-BIRA	Full lysis	16
NPM1_N-BIRA	IB	17
TF_N-BIRA	IB	16
YWHAZ_N-BIRA	IB	16

5.3.2 Purification

5.3.2.1 *Akta prime*

All batches were appeared to elute from the column successfully.

5.3.2.2 *Dialysis*

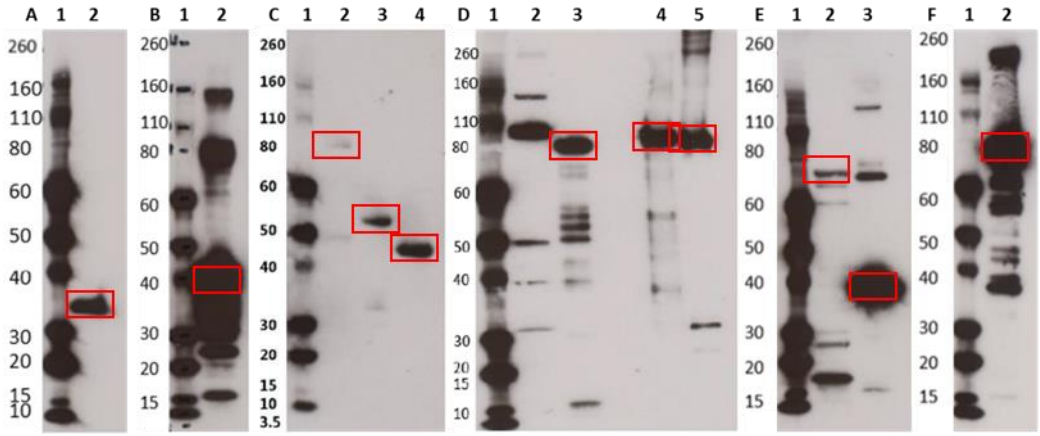
Dialysis/desalting proceeded without issue, with the exception of DDX3X_N-BirA. During the 2MDB to 1MDB step, this protein showed precipitation. The precipitated solution of DDX3X_N-BirA was centrifuged (13,000g, 30 minutes) and the supernatant subsequently filtered through 0.2µm filters. This process led to full removal of the precipitate. Comparison of yield prior to and after precipitate removal shows a reduction of yield from 43mgs to 2.2mgs.

5.3.2.3 *Quantification and characterisation*

Yield of each protein ranged from 1.5mg to 86.2mg (Table 21), at concentrations ranging from 100µg/ml to 2092µg/ml. Bands of the correct MW were present for every protein (Figure 29). Variation in yield is primarily dependent on the original production of the protein in *E. coli*. However, the purification process can strongly influence the final yield, as seen with the result of DDX3X_N-BirA, for which most 95% of the protein was lost due to precipitation.

Table 21 – Concentration and yields of the antigen batches produced

Protein	MW (kDa)	Concentration (µg/ml)	Yield (mg)
AIF1_N-BirA	31.6	157	71.4
CDKN1B_C-BirA	36.9	375	45
DDX3X_N-BirA	88	100	2.2 (precipitated)
HNRNPA2_C-BirA	35.9	1237	86.6
HNRNPL_C-BirA	77	496	36.5
HSPA4_N-BirA	87.4	125	61.2
HSPD1_N-BirA	72.9	797	23.9
MMP9_C-BirA	81.69	189	1.55
NPM1 _N-BIRA	47.5	1212	13.9
TF_N-BirA	91.9	121	55.5
YWHAZ_N-BirA	42.5	2092	67.9



Protein	MW (kDa)	Lane
AIF1_N-BirA	31.6	A2
CDKN1B_C-BirA	36.9	B2
DDX3X_N-BirA	88	C2
HNRNPA2_C-BirA	35.9	E3
HNRNPL_C-BirA	77	F2
HSPA4_N-BirA	87.4	D5
HSPD1_N-BirA	72.9	E2
MMP9_C-BirA	81.69	D2
NPM1 _N-BIRA	47.5	C3
TF_N-BirA	91.9	D4
YWHAZ_N-BirA	42.5	C4

Figure 29 – Anti-HIS western blots of each protein. Lane 1 of each blot contains Novex sharp prestained protein standard. Red boxes highlight bands of correct MW. Lane D2 contains a protein purified for a separate Oncimmune project.

5.4 Discussion

The purification scale-up was successful for all proteins, when compared with the HTP format but yield variation could be observed. For any proteins showing low yields, it may be necessary to perform further scale-up if they are to be used as part of a commercial assay, as they may not currently be cost effective. Unpublished data from other Oncimmune studies (see section 5.1) has shown cleaner signal for these FPLC purified proteins compared to HTP proteins, this can hopefully be seen in the section Chapter 6 as many of the same samples with be used that were also used in section Chapter 4, allowing signal differences can be compared.

Avoidance of precipitation was a key step in the purification process. Only DDX3X_N-BirA showed precipitation and the precipitate was successfully removed. Precipitation is caused by unstable insoluble proteins that are heavily aggregated. This can interfere with many quantification techniques which subsequently leads to inaccurate dilutions. Aggregation of the antigen can also block epitopes and reduce antibody binding.

Chapter 6. Combining leads

6.1 Introduction

Oncimmune have previously tested 40 AAbs for the detection of HCC[170]. These AAbs, except for NY-ESO-1, were different from those used in section Chapter 4. Some of the AAbs identified previously showed differentiation between samples from patients with HCC and NCCLD, whilst others did not. This section will focus on improving the results from Oncimmune's previous work by combining seven of the AAbs from previous work with 14 selected leads from the discovery assay. The sample set will be primarily the same as section Chapter 4, this will allow us to assess the results of antigen scale up as well as additivity. All batches of antigens used in this assay were therefore FPLC purified (section Chapter 5). The following antigens will be used in this experiment:

From section Chapter 4:

- HSPA4_N-BirA
- VIM_BirA
- EPCAM_BirA
- TF_N-BirA
- AIF1_N-BirA
- MMP9_C-BirA
- HNRNPL_C-BirA
- CDKN1B_C-BirA
- HSPD1_N-BirA
- HNRNPA2_C-BirA
- SALL4B_C-BirA
- NPM1_N-BirA
- YWHAZ_N-BirA
- DDX3X_N-BirA

From Oncimmune's previous work:

- NY-ESO-1_BirA
- CAGE_BirA
- MAGEA4_BirA
- P62_C-BirA

- RALA_C-BirA
- GBU4-5_BirA
- SOX2_BirA

6.2 Methods

6.2.1 Serum samples

The HCC and healthy control samples in this sample set are the same as those in the discovery assay described earlier (section 4.2.1). There was one fewer sample in the HCC group and four fewer in the healthy control group due to insufficient volume remaining. There were 22 sample changes to the NCCLD group due to insufficient volume, replacement samples were sourced from the QMC BRU. Sample demographics can be seen in Table 22-Table 25. As with the discovery assay, the NCCLD population has a lower mean age. This is partly attributed to the nature of HCC progressing from NCCLD, leading to the average age of HCC patients being higher than that of patients with NCCLD.

Table 22 - Age and gender data of the assay samples by group

Demographic factor	HCC patients	NCCLD patients	Individuals with no evidence of malignancy or liver disease
Number	98	99	95
Mean age	62.6	50.7	62.7
Age range	30-91	30-82	30-84
% Male	69	69	68

Table 23 - Primary tumour size of the HCC patients (where available)

	Number of patients with data available	Mean	Min-Max
Primary tumour size (cm)	86	6.02	0.4-19

Table 24 - HCC stage of the assay samples. N/A = stage data unavailable

TNM stage	Number
1	35
2	20
3	21
4	2
N/A	20

Table 25 - Aetiology of the NCCLD controls

NCCLD background	Number
Hepatitis C viral infection	58
Hepatitis B viral infection	31
Alcoholic Liver Disease	6
Autoimmune Hepatitis	2
Haemochromatosis	1
Primary Biliary Cirrhosis	1
Total	99

6.2.2 Plate coating

Seven antigens plus VOL were coated per plate were diluted to a semi log titration of 160nm, 50nm, 16nm, 5nm and 1.6nm concentrations. Each antigen point was plated in duplicate as seen in the plate layout (Figure 30-Figure 32). Antigen coating was performed as in section 4.2.3.

	0nM	1.6nM	5nM	16nM	50nM	160nM	0nM	1.6nM	5nM	16nM	50nM	160nM
	1	2	3	4	5	6	7	8	9	10	11	12
A	HNRNPL_C-BirA						NY-ESO-1_BIRA					
B	VOL						VIM_BirA					
C	MMP9_C-BirA						HSPA4_N-BirA					
D	AIF1_N-BirA						TF_N-BirA					
E	NY-ESO-1_BIRA						HNRNPL_C-BirA					
F	VIM_BirA						VOL					
G	HSPA4_N-BirA						MMP9_C-BirA					
H	TF_N-BirA						AIF1_N-BirA					

Figure 30 - Assay plate one layout

	0nM	1.6nM	5nM	16nM	50nM	160nM	0nM	1.6nM	5nM	16nM	50nM	160nM
	1	2	3	4	5	6	7	8	9	10	11	12
A	SALL4B_C-BIRA						EPCAM_BIRA					
B	VOL						CDKN1B_C-BIRA					
C	NPM1_N-BIRA						HSPD1_N-BIRA					
D	YWHAZ_N-BIRA						HNRNPA2_C-BIRA					
E	EPCAM_BIRA						SALL4B_C-BIRA					
F	CDKN1B_C-BIRA						VOL					
G	HSPD1_N-BIRA						NPM1_N-BIRA					
H	HNRNPA2_C-BIRA						YWHAZ_N-BIRA					

Figure 31 - Assay plate two layout

	0nM	1.6nM	5nM	16nM	50nM	160nM	0nM	1.6nM	5nM	16nM	50nM	160nM
	1	2	3	4	5	6	7	8	9	10	11	12
A	RALA_C-BIRA						DDX3X_N-BIRA					
B	VOL						CAGE_BIRA					
C	GBU4-5_BirA						MAGEA4_BIRA					
D	SOX2_BIRA						P62_C-BIRA					
E	DDX3X_N-BIRA						RALA_C-BIRA					
F	CAGE_BIRA						VOL					
G	MAGEA4_BIRA						GBU4-5_BirA					
H	P62_C-BIRA						SOX2_BIRA					

Figure 32 - Assay plate three layout

6.2.3 Assay protocol

The assay was performed as described in section 4.2.4

6.2.4 Data analysis

Data was analysed at antigen concentrations of 160nM and 50nM, whilst the other dilutions could be used for data analysis, to be consistent with the previous discovery assay only the 160nM and 50nM were used. The other data points are useful as quality control checks as the dilution curves for each sample can be analysed to look for issues in antigen dilution, plate coating and non-specific binding/anomalous wells. For individual marker analysis, cut-offs were chosen as the value which gave the highest Youden's J statistic for HCC vs

NCCLD patients where the specificity was >=95% and the J statistic was >1. Significance analysis for difference sensitivity of the panel at individual stages of tumour was assessed using McNemar's test. The Stata (Stata corp.) package roccomp was used to assess AUC improvements between tests. Panel association with stage was analysed using logistic regression. Bland-Altman plots were created to analyse the effect of batch reproduction on assay repeatability.

6.3 Results

6.3.1 Individual markers

The performance of each marker can be seen in Table 26. Only AAbs which satisfied the criteria set out in the methods have been listed in the table below. Of the 21 AAbs tested, 17 could differentiate HCC from NCCLD patients. A selection of markers are represented as dot plots in Figure 33.

Table 26 – Number of positives and calculated sensitivity and specificity of each marker by sample group. N/A implies that no cut-off could be assigned which matched the criteria.

	HCC		NCCLD		Healthy controls	
Marker	**Positives**	**Sensitivity (%)**	**Positives**	**Specificity (%)**	**Positives**	**Specificity (%)**
AIF1_N-BirA 160nM	1	1.02	0	100.00	0	100.00
AIF1_N-BirA 50nM	1	1.02	0	100.00	0	100.00
CAGE_BirA 160nM	9	9.18	2	97.98	6	93.68
CAGE_BirA 50nM	12	12.24	3	96.97	9	90.53
CDKN1B_C-BirA 160nM	N/A	N/A	N/A	N/A	N/A	N/A
CDKN1B_C-BirA 50nM	N/A	N/A	N/A	N/A	N/A	N/A
DDX3X_N-BirA 160nM	5	5.10	2	97.98	2	97.89
DDX3X_N-BirA 50nM	3	3.06	2	97.98	1	98.95
EPCAM_BirA 160nM	4	4.08	2	97.98	5	94.74
EPCAM_BirA 50nM	3	3.06	0	100.00	3	96.84
HNRNPA2_C-BirA 160nM	6	6.12	4	95.96	6	93.68
HNRNPA2_C-BirA 50nM	N/A	N/A	N/A	N/A	N/A	N/A
HSPA4_N-BirA 160nM	4	4.08	1	98.99	0	100.00
HSPA4_N-BirA 50nM	4	4.08	1	98.99	0	100.00
HSPD1_N-BirA 160nM	N/A	N/A	N/A	N/A	N/A	N/A
HSPD1_N-BirA 50nM	2	2.04	2	97.98	1	98.95

MAGEA4_ BirA 160nM	3	3.06	0	100.00	0	100.00
MAGEA4_ BirA 50nM	4	4.08	1	98.99	0	100.00
MMP9_C-BirA 160nM	2	2.04	1	98.99	0	100.00
MMP9_C-BirA 50nM	5	5.10	4	95.96	5	94.74
NY-ESO-1_BirA 160nM	12	12.24	4	95.96	14	85.26
NY-ESO-1_BirA 50nM	8	8.16	0	100.00	4	95.79
P62_C-BirA 160nM	5	5.10	2	97.98	1	98.95
P62_C-BirA 50nM	2	2.04	0	100.00	0	100.00
RALA_C-BirA 160nM	3	3.06	1	98.99	2	97.89
RALA_C-BirA 50nM	4	4.08	1	98.99	4	95.79
SALL4B_C-BirA 160nM	6	6.12	3	96.97	1	98.95
SALL4B_C-BirA 50nM	5	5.10	3	96.97	2	97.89
SOX2_BirA 160nM	3	3.06	0	100.00	1	98.95
SOX2_BirA 50nM	5	5.10	1	98.99	4	95.79
TF_N-BirA 160nM	6	6.12	4	95.96	1	98.95
TF_N-BirA 50nM	4	4.08	3	96.97	1	98.95
YWHAZ_N-BirA 160nM	6	6.12	4	95.96	5	94.74
YWHAZ_N-BirA 50nM	N/A	N/A	N/A	N/A	N/A	N/A

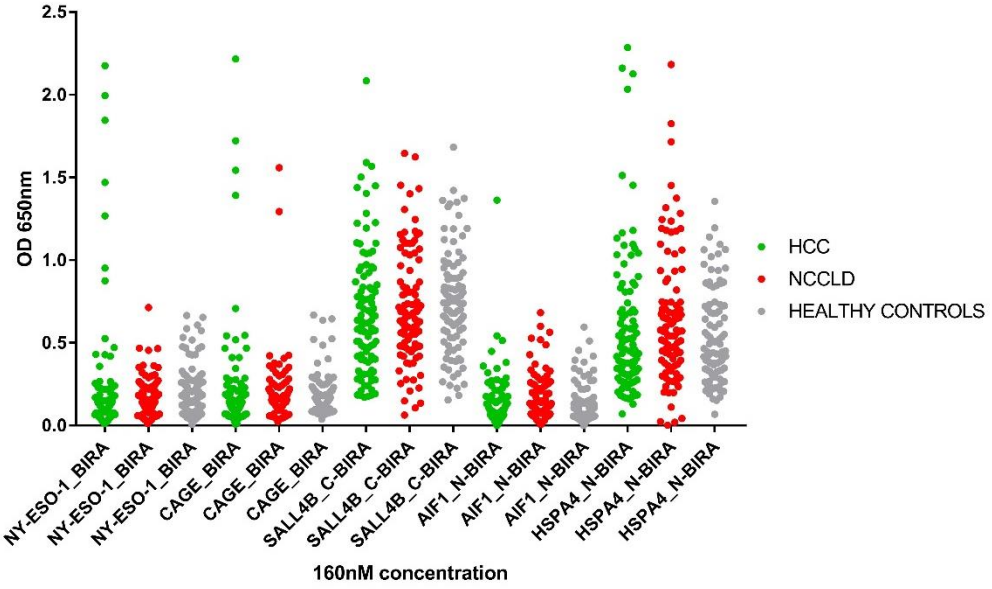


Figure 33 - Dot plot results of a selection of AAbs run in this study split by sample type

6.3.2 Net reclassification improvement analysis

Individual AAb performance remains low as seen in section Chapter 4. AAbs were analysed by net reclassification improvement (NRI) to create a panel in which, a single positive result for any AAb of the panel would give a positive result (indicating HCC). Starting with the AAb with the highest Youden's J statistic (CAGE, 0.09), AAbs were added one at a time until no AAb would improve the accuracy of the panel (Table 27). Overall sensitivity of the panel was 36.7% sensitivity for HCC at 89.9% specificity in the NCCLD. This equates to an increase of 24.5% in sensitivity for a drop of only 7% in specificity compared to CAGE alone and a significant increase in AUC from 0.546 to 0.633 (p=0.0006).

Table 27 - NRI formation of the AAb panel, sensitivity and specificity values shown are cumulative as each marker is added to the panel

Addition order	Model	NRI score	Sensitivity (%)	Specificity NCCLD (%)	Specificity healthy controls (%)
1	CAGE_BirA 50nM	0.092449	12.24	96.97	90.53
2	NY-ESO-1_BirA 50nM	0.05102	17.35	96.97	89.47
3	RalA_C- BirA 50nM	0.030816	21.43	95.96	88.42
4	EpCAM_BirA 50nM	0.030612	24.49	95.96	88.42
5	SOX2_ BirA 160nM	0.020408	26.53	95.96	87.37
6	AIF1_N-BirA 160nM	0.010204	27.55	95.96	87.37
7	MMP9_C-BirA 160nM	0.010408	29.59	94.95	87.37
8	MAGE-A4_BirA 160nM	0.010204	30.61	94.95	87.37
9	MAGE-A4_ BirA 50nM	0.000204	31.63	93.94	87.37
10	SALL4B_C-BirA 160nM	0.010612	34.69	91.92	86.32
11	DDX3X_N- BirA 160nM	0.000408	36.73	89.90	84.21

6.3.3 Addition of AFP

AFP values were taken from the discovery assays in section 4. For the few samples changed values were obtained via the same method as section 4. Due to this, performance of AFP alone was as seen in section 4. It is still of interest to see if AFP and the AAb panel identified in section 6.3.2 are still additive to overall sensitivity, as four of the ten AAbs in this panel were not used in the discovery panel identified in section 4. AFP had a sensitivity of 31.6% in the HCC samples and specificities of 100% in both control groups. Addition of AFP to the AAb panel raised sensitivity of the panel from 36.7% to 55.1% without reducing the specificity in either control group. This gave a statistically significant increase in AUC from 0.634 to 0.725 ($p<0.0001$).

6.3.4 Performance by stage

The sensitivity of the AAb panel, AFP and a combination of both by stage are shown in Table 28. The AAb panel positivity was not significantly affected by

stage (p=0.36). AFP positivity was significantly associated with an increase in stage (p=0.016). The combination of the AAb panel and AFP was significantly associated with increasing stage (p=0.022). Although addition of AFP to the AAb panel lead to an association with higher stage, sensitivity in stage one HCCs was higher than either the AAb panel (p=0.0455) or AFP (p=0.0047) alone. The combined panel was also significantly better at detecting stage two HCCs than AFP alone (p=0.0455) but not the AAb panel (p=0.0833).

Table 28 - Performance of the AAb panel, AFP and a combination of both by HCC TNM stage

TNM Stage	Positives	Negatives	Sensitivity
AAb Panel			
1	10	25	29
2	8	12	40
3	8	13	38
4	1	1	50
N/A	9	11	45
AFP			
1	6	29	17
2	7	13	35
3	11	10	52
4	1	1	50
N/A	9	11	45
AAb Panel + AFP			
1	14	21	40
2	11	9	55
3	14	7	67
4	2	0	100
N/A	13	7	65

6.3.5 **Assay reproducibility**

Bland-Altman plots were created for each antigen/concentration for all samples which overlapped between the discovery assay and this assay. Table 29 shows the mean difference, upper and lower 95% confidence intervals and the 95% confidence interval range for each antigen/concentration. Narrower confidence inter ranges imply closer agreement in values between assays. The top six ranked antigens/concentrations are both concentrations of VIM, NY-ESO-1 and EPCAM. All three of these used the same batch of antigen between this assay and the discovery assay so indicate good reproducibility between

assay runs. The negative mean difference for each batch indicate that signals were lower in this assay than the discovery assay, which is expected as increase in purity of antigen should lead to lower non-specific binding, hence lower mean signals in this assay compared to the discovery assay. Figure 34 shows example Bland-Altman plots depicting three scenarios of antigen reproduction. Plot A shows NY-ESO-1 at 50nM, for this, the same batch of antigen was used in each assay, confidence intervals are tight and samples which appear to be true positive results lie close to a difference in signal of 0. Plot B shows HSPA4 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however samples which appear to be true positives have a high mean signal and difference in signal between the two batches close to 0. Plot C shows NPM1 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however all signal appeared to be from non-specific binding as there are no samples with a high mean signal and difference in signal between the two batches close to 0.

Table 29 - Parameters established from Bland-Altman plots for each antigen/concentration ranked by the 95% confidence interval range

Antigen	Conc. (nM)	Mean difference	95% Lower limit	95% Upper limit	95% CI Range
VIM	50	-0.019	-0.128	0.091	0.218
VIM	160	-0.030	-0.192	0.132	0.324
NY-ESO-1	50	-0.029	-0.203	0.144	0.348
EPCAM	50	-0.023	-0.211	0.165	0.376
NY-ESO-1	160	-0.033	-0.249	0.184	0.433
EPCAM	160	0.019	-0.202	0.241	0.443
CDKN1B	160	-0.016	-0.367	0.335	0.703
AIF1	160	-0.128	-0.497	0.242	0.739
TF	160	-0.005	-0.384	0.375	0.759
TF	50	-0.020	-0.406	0.366	0.772
CDKN1B	50	-0.270	-0.672	0.132	0.805
MMP9	160	-0.044	-0.466	0.379	0.846
NPM1	160	-0.300	-0.724	0.124	0.848
HSPA4	160	-0.069	-0.517	0.380	0.898
SALL4B	50	-0.001	-0.454	0.451	0.905
HNRNPA2	160	-0.184	-0.662	0.294	0.956
AIF1	50	-0.363	-0.854	0.129	0.983
NPM1	50	-0.368	-0.868	0.133	1.001
HNRNPL	160	-0.169	-0.676	0.339	1.014
HSPD1	50	-0.412	-0.926	0.102	1.028
SALL4B	160	-0.030	-0.552	0.492	1.043
HNRNPL	50	-0.149	-0.696	0.398	1.094
HNRNPA2	50	-0.436	-0.996	0.124	1.120
MMP9	50	-0.514	-1.075	0.047	1.122
HSPD1	160	-0.406	-0.970	0.158	1.128
HSPA4	50	-0.618	-1.231	-0.005	1.227
YWHAZ	160	-0.413	-1.041	0.216	1.258
YWHAZ	50	-0.714	-1.466	0.038	1.504
DDX3X	50	-0.511	-1.274	0.253	1.526
DDX3X	160	-0.769	-1.595	0.057	1.652

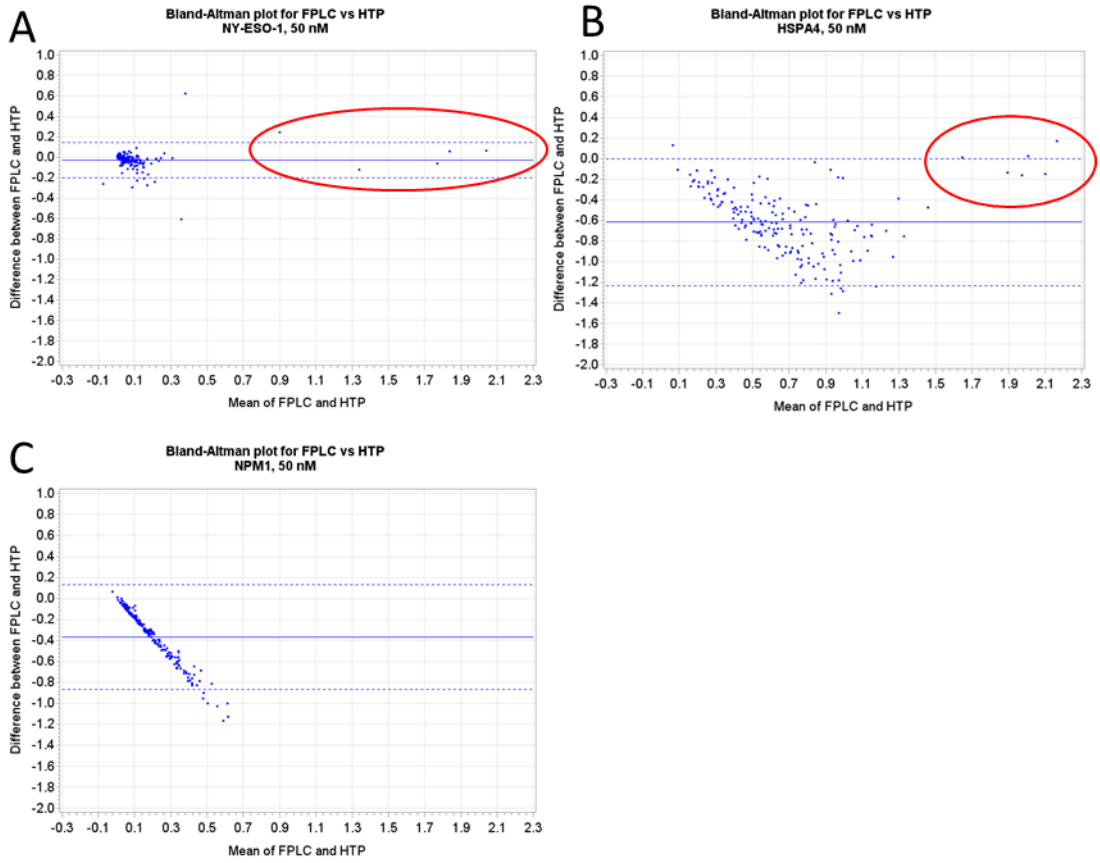


Figure 34 – Example Bland-Altman plots showing 3 situations of antigen reproducibility between this assay and the discovery assay. Plot A shows NY-ESO-1 at 50nM, for this, the same batch of antigen was used in each assay, confidence intervals are tight and samples which appear to be true positives (red circle) results lie close to a difference in signal of 0. Plot B shows HSPA4 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however samples which appear to be true positives (red circle) have a high mean signal and difference in signal between the two batches close to 0. Plot C shows NPM1 at 50nM, for this, switching to a FPLC antigen batch leads to a negative trend in difference indicating lower signal in this assay compared to the discovery, however all signal appeared to be from non-specific binding as there are no samples with a high mean signal and difference between the two batches close to 0.

6.4 Discussion

As seen in previous sections, no single AAb could distinguish HCC from NCCLD patients to a level that would be clinically useful. For some markers, only one of the two concentrations (160nM or 50nM) was useful. This could imply a problem with concentration estimation, if the estimated concentration of the antigen is too high, there will be too little antigen on the plate to give a measurable signal range, whereas if the antigen concentration is estimated too low, there will be too much antigen on the plate, pushing signals above the max range of the assay and increasing levels of non-specific binding.

Combining the results of individual AAbs into a panel significantly improved the accuray from 0.546 to 0.633 over the AAb (CAGE) with the highest AUC alone. This panel of 10 AAbs identified by NRI could distinguish HCC patients of all stages without significant association with early or late stage tumours. Positivity of AFP, the most commonly used blood-based marker for HCC, was significantly associated with increasing tumour stage. The chance of a positive prognostic outcome decreases as tumour stage increases. This is commonly seen for AFP as the amount secreted often correlated with tumour size[256]. Addition of AFP to the AAb panel significantly increased the overall AUC from 0.633 and 0.658 respectively to 0.725. Sensitivity of the combination of AFP and the AAb panel was with significantly associated with increasing stage, however was still better at detecting stage one HCCs (40%) than either the AAb (27%) panel or AFP alone (17%). So far, all this work has been undertaken using the same set of samples. This leaves the results susceptible to over fitting. Assessment of this AAb panel, and the effect of combination with AFP, should be evaluated in an independent sample set.

Chapter 7. Training and validation

7.1 Introduction

The previous experiments demonstrated the ability of AAbs to distinguish HCC patients from those with non-cancerous liver disease and healthy control patients. After discovery (section Chapter 4) and confirmation assays (section Chapter 6), 10 lead AAbs were identified. In both experiments, the selection of an AAbs into a panel was more accurate than any AAb alone and was also additive to circulating AFP antigen measurement. The previous assays were run on primarily the same sample set, so it is hard to assess overfitting of the data. Assays in this section measured AAbs against the lead antigens identified from previous section and run them against a new set of samples. This sample set was split into training and validation. The training set was used to set cut-offs for the AAbs, the performance of which was then validated in the validation set. This was performed to identify what level of overfitting, if any, is present in panel formation. As there were 14 spaces for antigens on this plate layout, the next two best AAbs from Oncimmune's previous work[251], which were not tested in section Chapter 6 (CYCLINB1 and AFP), and two AAbs from section Chapter 6 which performed well individually but did not make it into the panel (HSPA4_N-BirA and TF_N-BirA), were also included. The 14 antigens used to measure AAbs were: CAGE_BirA, NY-ESO-1_BirA, MAGEA4_BirA, SOX2_BirA, TF_N-BirA, EPCAM_N-BirA, HSPA4_N-BirA, SALL4B_C-BirA, AIF1_N-BirA, DDX3X_N-BirA, MMP9_N-BirA, CYCLINB1_C-BirA, RALA_C-BirA and AFP_C-BirA. MMP9 was cloned into the N-BirA vector by Oncimmune for use in this assay for commercial purposes.

7.2 Methods

7.2.1 Serum samples

Serum samples for this study were collected from four separate locations,

1. Egypt – A prospective collection of samples from three groups, patients with HCC, patients with NCCLD and healthy controls

2. STOP-HCV – Patients with HCV related HCC, patients with HCV related cirrhosis, patients with HCV but no evidence cirrhosis

3. PAS – A collection of healthy controls collected by Oncimmune

4. QMC BRU – Samples collected by the Queens Medical Centre biomedical research unit. Patients with HCC or NCCLD

These samples were then split into a training set of: 200 HCC patients, 200 NCCLD and 200 healthy controls (Table 30) and a validation set of: 100 HCC patients, 100 NCCLD and 100 healthy controls (Table 31).

Table 30 - Training set sample origins

Sample origin	HCC	NCCLD	Healthy
Egypt	159	52	71
STOP-HCV	29	125	0
BRU	12	23	0
PAS	0	0	129

Table 31 - Validation set sample origins

Sample origin	HCC	NCCLD	Healthy
Egypt	84	41	30
STOP-HCV	10	58	0
BRU	6	1	0
PAS	0	0	70

7.2.2 Plate coating

Plate coating was performed as described in section 6.2.2

7.2.3 Assay protocol

Assay was performed as described in section 4.2.4

7.2.4 AFP measurement

Circulating AFP was measured in the serum samples using a commercially available ELISA (Monobind) following the manufacturer's instructions. A commonly-used cut-off of 200ng/ml was applied to assess positivity.

7.2.5 Data analysis

7.2.5.1 *Quality control*

Samples were excluded from the analysis if they exhibited high background or non-specific binding. Non-specific binders were determined by excluding samples where the mean VOL (see section 4.2.7 for further details) corrected OD of all 160nM wells was >1.2OD. High background signal was determined by excluding any samples where the uncorrected mean OD of all 0nM wells was >1.2.

7.2.5.2 *Calibration*

To correct for inter-assay variation, ODs were converted to relative units (RU) using a calibrator fluid. Calibrator fluids were identified by Oncimmune as showing specific and reproducible binding to each antigen in the assay. An 8-point dilution curve was formed for each antigen at 160nM and 50nM. A standard four-parameter logistic model was fit for each antigen and dilution. The RU value was then calculated as follows:

$$RU = f - \frac{\ln\left(\frac{m-c}{OD-c} - 1\right)}{b} + logpot + \ln(1000)$$

Where c=minimum asymptote, m=maximum asymptote, b=slope factor, f=ec50, logpot= log(relative potency)

7.2.5.3 *Cut-offs and panel formation*

Panel formation was accomplished using a modified Monte-Carlo analysis[257–259] to randomly pick cut-offs for each AAb and produce the overall panel performance in terms of sensitivity, specificity and Youden's score. Due to the sheer number of outcomes, 400 cut-offs for each of 14 AAbs at 2 dilutions, restrictions were applied to which cut-offs could be picked. Only cut-offs which gave an individual specificity of >=95% were included as they are more likely to form a panel with the high specificity needed for a screening test. The cut-off set giving the highest Youden's score with a specificity>=85% will be picked and applied to the validation set. Significance analysis for difference sensitivity of the panel at individual stages of tumour was assessed using McNemar's test.

The Stata (Stata corp.) package roccomp was used to assess accuracy improvements between tests. Panel association with stage was analysed using logistic regression.

7.3 Results

7.3.1 Training

7.3.1.1 *Quality control*

Due to failures of the QC limits, 19 samples were excluded from the training set analysis. 12 of these samples were from the HCC group, 4 from the NCCLD group and 3 from the healthy controls. All samples were excluded based on high non-specific binding.

7.3.1.2 *Monte Carlo simulation*

A Monte-Carlo simulation was run for 100 million iterations, and the training panels ranked via the Youden's score obtained from the sensitivity combined with the specificity of the NCCLD samples. A number of 5300 panels satisfied the constraints. Of the 5300 panels, the best performance (highest Youden's score) showed a sensitivity of 46.27% at a specificity of 88.27% in the NCCLD group and 86.8% in the healthy controls. This panel consisted of 11 AAbs with a cut-off assigned at 160nM and 50nM concentrations for all AAbs except for DDX3X_N-BirA and HSPA4_N-BirA, which were only included at 160nM. MMP9_N-BirA was only included at 50nM. RALA_C-BirA was only included at 160nM due to a contamination issue with the 50nM dilution identified during the assay and therefore excluded from the analysis. Performances for each AAb at the cut-off assigned to the panel are shown in Table 32. TF_N-BirA at 160nM had 0% sensitivity and 99.5% specificity, so was adding nothing to the performance of the panel and is an artefact of the Monte-Carlo method. The panel was significantly better at distinguishing HCC from the NCCLD controls than the AAb with the highest Youden's score (HSPA4_N-BirA) ($p<0.0001$).

Table 32 - Performance of the individual AAbs at the cut-offs selected by the Monte Carlo direct search in the training set

AAb	Sensitivity (%)	Specificity NCCLD (%)	Specificity healthy controls (%)
AFP_C-BIRA 160nM	4.26	100.00	98.50
AFP_C-BIRA 50nM	4.26	99.49	99.50
CAGE_BIRA 160nM	12.77	100.00	99.50
CAGE_BIRA 50nM	14.89	97.96	99.00
CYCLINB1_C-BIRA 160nM	1.06	97.96	99.50
CYCLINB1_C-BIRA 50nM	0.53	98.98	99.50
DDX3X_N-BIRA 160nM	5.32	96.94	97.50
EPCAM_BIRA 160nM	2.13	97.45	99.00
EPCAM_BIRA 50nM	3.19	97.96	98.50
HSPA4_N-BIRA 160nM	27.66	89.80	92.90
MAGEA4_BIRA 160nM	6.91	98.47	99.50
MAGEA4_BIRA 50nM	4.79	100.00	100.00
MMP9_N-BIRA 50nM	3.19	99.49	100.00
NY-ESO-1_BIRA 160nM	9.04	96.43	98.00
NY-ESO-1_BIRA 50nM	11.17	95.41	98.00
RALA_C-BIRA 160nM	1.60	99.49	99.00
TF_N-BIRA 160nM	0.00	99.49	99.50
TF_N-BIRA 50nM	2.13	99.49	98.00

Stage data was available for 153 of the 188 HCC samples. The sensitivity of the panel by stage is shown in Table 33. Stage was not a significant predictor of the panel positivity (odds ratio=1.211, p=0.333).

Table 33 - Performance of the AAb panel by BCLC stage in the training set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	1	0.0
BCLC A	20	18	52.6
BCLC B	35	41	46.1
BCLC C	18	12	60.0
BCLC D	5	3	62.5

7.3.1.3 *AFP measurement*

AFP antigen measurement showed a sensitivity of 28.2% for HCC samples at a specificity of 99.5% for both NCCLD and healthy samples. The sensitivity by

stage is shown in Table 34. Sensitivity of AFP was significantly associated with increasing stage (Odds ratio=2.894, p<0.001).

Table 34 - Performance of AFP (>200ng/ml) by BCLC stage in the training set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	1	0.0
BCLC A	3	35	7.9
BCLC B	25	51	32.9
BCLC C	17	13	56.7
BCLC D	5	3	62.5

7.3.1.4 *AAB + AFP*

Combined performance of AFP and the AAb panel had a sensitivity of 58.5% and specificities of 87.2% and 86.3% for the NCCLD and healthy controls respectively. Addition of AFP to the panel significantly increased the AUC when distinguishing HCC samples from NCCLD (0.670 to 0.729, p<0.0001) or healthy controls (0.665 to 0.724, p<0.0001).

The sensitivity by stage of the combined panel is shown in Table 35. Stage data was available for 153 of the 188 HCC samples. Stage was not a significant predictor of the panel positivity (odds ratio=1.211, p=0.333). Addition of AFP had no effect on the detection of BCLC 0, A or D cancers but did significantly increase the detection of stage B (p=0.001) and C (p=0.0078) tumours.

Table 35- Performance of the AAb panel in combination with AFP by BCLC stage in the training set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	1	0.0
BCLC A	20	18	52.6
BCLC B	35	41	46.1
BCLC C	18	12	60.0
BCLC D	5	3	62.5

7.3.2 **Validation**

7.3.2.1 *QC*

Due to failures of the QC limits, 2 HCC samples were excluded from the validation set analysis leaving a final sample set of 98 HCC, 100 NCCLD and 100

healthy controls. Both HCC samples were excluded as due to high non-specific binding.

7.3.2.2 *AAb panel*

Cut-offs determined from the training set were applied to the validation samples resulting in a sensitivity of 33.7% and a specificities of 89% and 95% for the NCCLD and healthy controls respectively. Individual AAb performances are shown in Table 36. There was no significant difference in the ability of the AAb panel to distinguish HCC from NCCLD between the training and validation set (p=0.1143).

Table 36 - Performance of the individual AAbs in the validation set at the cut-offs selected in the training set

AAb	Sensitivity (%)	Specificity NCCLD (%)	Specificity healthy controls (%)
AFP_C-BirA 160nM	0.00	100.00	99.00
AFP_C-BirA 50nM	0.00	99.00	99.00
CAGE_BirA 160nM	8.16	100.00	100.00
CAGE_BirA 50nM	14.30	98.00	99.00
CYCLINB1_C-BirA 160nM	1.02	99.00	99.00
CYCLINB1_C-BirA 50nM	0.00	100.00	100.00
DDX3X_N-BirA 160nM	0.00	100.00	100.00
EPCAM_BirA 160nM	3.06	100.00	99.00
EPCAM_BirA 50nM	10.20	99.00	98.00
HSPA4_N-BirA 160nM	5.10	98.00	98.00
MAGEA4_BirA 160nM	12.20	96.00	99.00
MAGEA4_BirA 50nM	10.20	96.00	97.00
MMP9_N-BirA 50nM	1.02	100.00	100.00
NY-ESO-1_BirA 160nM	2.04	100.00	99.00
NY-ESO-1_BirA 50nM	4.08	99.00	100.00
RALA_C-BirA 160nM	0.00	100.00	100.00
TF_N-BirA 160nM	0.00	100.00	100.00
TF_N-BirA 50nM	0.00	100.00	100.00

Stage data was available for 77 of the 98 HCC samples. The sensitivity of the panel by stage is shown in Table 37. Stage was not a significant predictor of the panel positivity (odds ratio=1.367, p=0.354).

Table 37 - Performance of the AAb panel by BCLC stage in the validation set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	0	N/A
BCLC A	7	10	41.2
BCLC B	13	28	31.7
BCLC C	9	9	50.0
BCLC D	1	0	100.0

7.3.2.3 *AFP measurement*

AFP antigen measurement had a sensitivity of 26.5% for HCC samples and specificities of 100% in the NCCLD and healthy controls. There was no significant difference in the ability to differentiate HCC samples from NCCLD controls between training and validation (p=0.8367). The sensitivity for each stage of tumour is shown in Table 38. Stage was not a significant predictor of the panel positivity (odds ratio=1.31, p=0.448).

Table 38 - Performance of AFP by BCLC stage in the validation set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	0	N/A
BCLC A	4	14	22.2
BCLC B	11	30	26.8
BCLC C	7	11	38.9
BCLC D	0	1	0.0

7.3.2.4 *AAB + AFP*

Combined performance of AFP and the AAb panel had a sensitivity of 52% and a specificities of 89% and 95% in the NCCLD and healthy controls respectively. Performance for HCC vs NCCLD controls was not significantly different from the training set (p=0.5223).

Stage data was available for 77 of the 98 HCC samples. The sensitivity of the panel for each tumour stage is show in Table 39. Stage was not a significant predictor of the panel positivity (odds ratio=1.587, p=0.169). Addition of AFP had no effect of detection of BCLC stage D tumours and had no significant

increase in BCLC A (p=0.5) or C tumours (p=0.125). Addition of AFP did significantly increase detection of BCLC B (p=0.0078) tumours.

Table 39 - Performance of the AAb panel in combination with AFP by BCLC stage in the training set

Stage	Positives	Negatives	Sensitivity (%)
BCLC 0	0	0	0.0
BCLC A	9	8	52.9
BCLC B	21	20	51.2
BCLC C	13	5	72.2
BCLC D	1	0	100.0

7.4 Discussion

The results from this assay confirm what has been seen throughout this thesis, no single AAb is capable of separating HCC from NCCLD to a level that would be clinically useful. However, as seen previously, a panel of multiple AAbs can significantly increase performance over one alone. This study also helps to answer the question of overfitting posed earlier. Whilst there is a difference in performance between the training set and validation set, there is not a significant difference in panel performance. Some AAbs perform better than others in the training set but vice-versa in the validation set, for example AFP_C-BirA at 160nM had a Youden's score of 0.043 in the training set but a score of 0 in validation whereas, MAGEA4_BirA at 50nM showed an increase in Youden's score from 0.048 to 0.062 from training to validation. This highlights the high degree of heterogeneity in cancer, especially HCC. For this reason, a panel is ultimately more useful than a single marker. There is some overlap in positivity between AAbs in the training set which may have been removed by other panel selection algorithms. EPCAM_BirA, for example, had very mediocre performance in the training set but much higher performance in the validation set. It may well be best to leave this redundancy in the panel to allow for cohort variation. This highlights the need to thoroughly test markers before considering them for clinical use. TF_N-BirA at 160nM did not contribute to either panel and had less than 100% specificity so this should be removed from the panel as there is no evidence of its clinical use and removing it has no effect on the performance of the panel in either training or validation. Whilst TF_N-

BirA 160nM does not contribute, TF_N-BirA 50nM does. This demonstrates the reason for testing multiple concentrations. The Bradford assay used for protein quantification in this thesis is not 100% accurate as it measures total protein and therefore does not allow for contamination with bacterial proteins which may be present at varying concentrations, or degradation of the target protein to forms which lose their ability to bind antibodies. In the case of TF_N-BirA, 160nM appears to contain too much protein, therefore condensing signals at the higher end and resulting in loss of resolution between results.

As with previous panels in this thesis, AFP was significantly additive to the AAb panel and the combination of both was significantly better than either alone increasing the AUC to 0.721 from 0.636 and 0.651 for AFP and the AAb panel respectively. There was no significant increase in detection of BCLC A cancers with the addition of AFP compared to the AAb panel alone, however this is unsurprising as AFP classically has poor performance in earlier stage HCCs. The performance in BCLC B cancers was significantly increased. Whilst BCLC 0/A tumours are the goal for detection due to high success rates of resection, BCLC B tumours still have a better prognosis than C and D. BCLC tumours are typically treated with TACE and can be resected or transplanted if this leads to tumour shrinkage and downgrading of stage[260]. Despite not increasing stage 0/A detection, AFP significantly adds to the panel without affecting the specificity so is a useful addition.

Another key takeaway from this experiment is the rate of un-assessable samples. This is much higher in the HCC group than the NCCLD group or healthy controls (14 vs 4 vs 3 respectively). All of these were ruled out due to apparent non-specific binding. The most likely cause of this non-specific binding is binding to the BirA tag added to each antigen. The BirA tag is a slightly modified bacterial protein present in many different strains. It has been shown that liver damage can cause an increase in bacterial translocation and may play a critical role in the progression of HCC[261]. This poses a question of whether these results could be included and that increased anti-bacterial antibodies could be a

marker of HCC, however this is likely to result in specificity issues with patients that are prone to bacterial infections and requires further research.

Chapter 8. Preliminary serial samples assay

8.1 Intro

As mentioned in section 1.6, patient specific profiling shows potential to improve on current biomarker performance. ROCA, an algorithm for longitudinal measurement of CA-125 has previously been validated for ovarian cancer screening[175]. Although this applies to an antigen rather than AAbs, data from Oncimmuine's previous experiments point towards similarities. In patients without cancer, AAb levels show very little variation, even allowing to identification of different samples from the same patient from a blinded data set (unpublished data). This should hypothetically allow the monitoring of patients for an increase in AAb levels past their normal limits. During early cancer development this difference may be noticeable, but not pass the traditional test cut-offs. To test this hypothesis, sequential samples from five high risk HCV patients who went on to develop HCC were analysed for their AAb responses.

8.2 Methods

8.2.1 Antigen selection

Due to limited sample availability only one assay plate could be run per sample (seven antigens). As this assay was performed early in the project, new antigens produced for this project were not ready. Therefore, seven antigens previously produced and tested by Oncimmune were used. These were picked on prior evidence of cancer/control differentiation in HCC. Thus NY-ESO-1_BirA, P62_C-BirA, CK8_BirA, GBU4-5_BirA, CYCLINB1_C-BirA, IMP1_C-BirA and HRAS2_C-BirA were chosen for the ELISA.

8.2.2 Autoantibody ELISA

Antigen dilution, plate coating and ELISA were performed as detailed in sections, 6.2.2 and 4.2.4.

8.3 Results

The full set of serial samples from the 5 HCC patients were run for the chosen 7 antigens at a full titration of 160nM, 50nM, 16nM, 5nM, and 1.6nM. No non-

specific binding was identified, so the 160nm results were analysed (Figure 35-Figure 39). Patient one shows a sharp rise in antibodies to CYCLINB1 and GBU4-5, as well as smaller rises in all other markers between sample three and four (Figure 35). Theses samples are both pre-diagnosis, however this patient went on to develop HCC post sampling. These rises could be indicative of HCC occurrence. Patient two also shows sharp rises in antibodies to CYCLINB1, GBU4-5 and HRAS2 between samples three and four. HCC was diagnosed at the time of sample three in this patient, with the initial sharp rise in CYCLINB1 AAbs occurring ten months prior (Figure 36). Patient three shows an increase in IMP1 AAbs between samples two and three which is then sustained 23 months later in the diagnosis sample (Figure 37). Patient four shows a rise in NY-ESO-1, GBU4-5, HRAS2 and CYCLINb1 AAbs within a month with only a slight decrease six months later (Figure 38). Patient five appears to have no change that could be predictive of HCC (Figure 39).

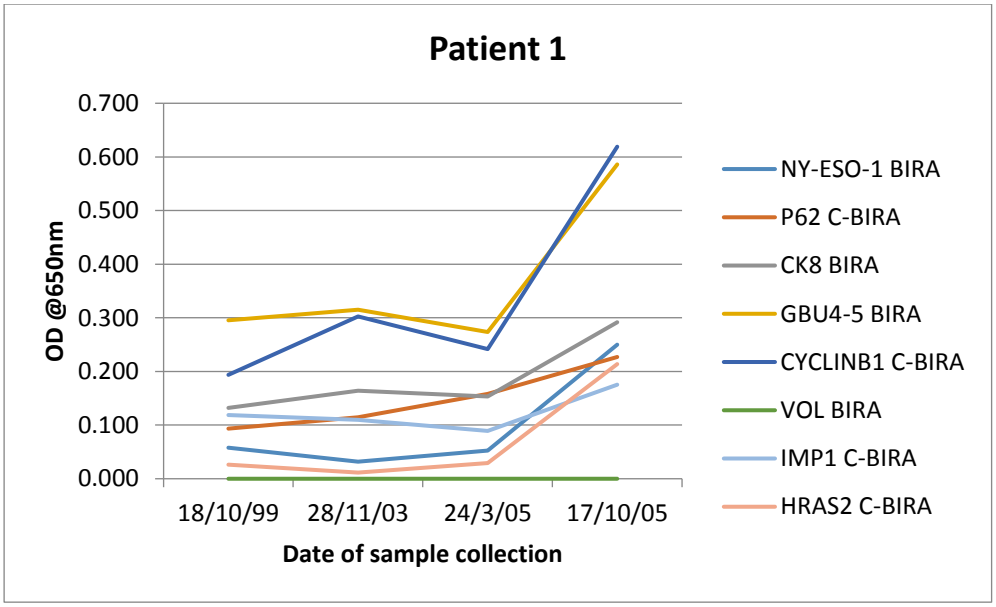


Figure 35 - OD readings at a protein concentration on 160nm of multiple samples from Patient 1. Patient 1 is male and 46 years old at the date of first sample collection. HCC was clinically after the date of final sample collection.

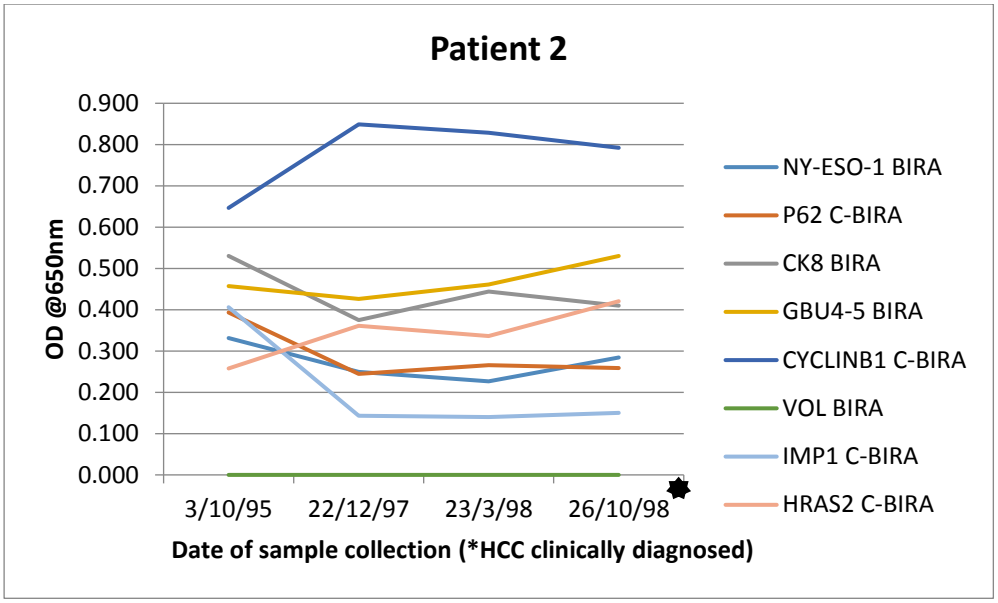


Figure 36 - OD readings at a protein concentration on 160nm of multiple samples from Patient 2. Patient 2 is female and 63 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken.

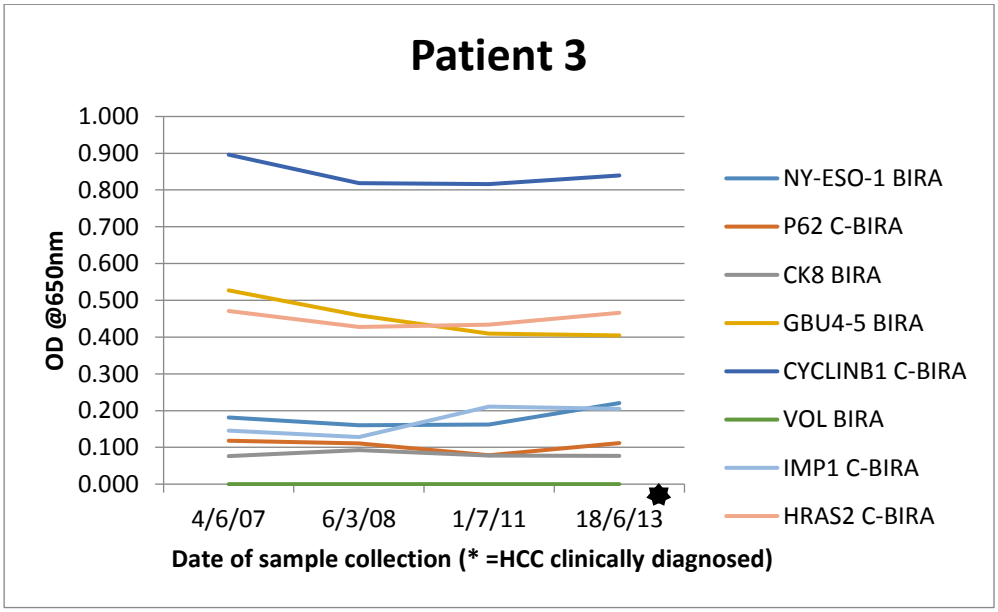


Figure 37 - OD readings at a protein concentration on 160nm of multiple samples from Patient 3. Patient 3 is female and 43 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken.

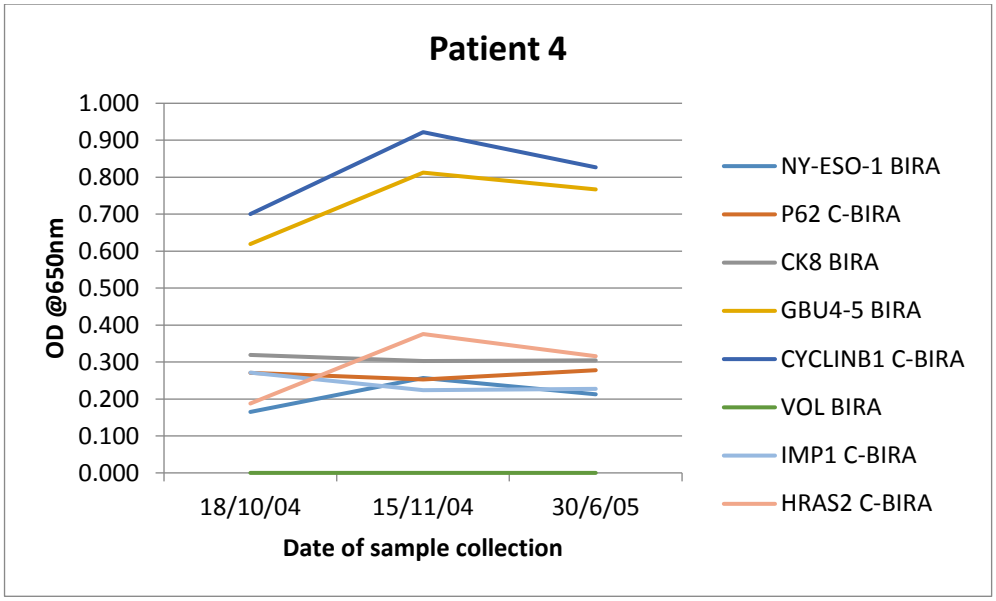


Figure 38 - OD readings at a protein concentration on 160nm of multiple samples from Patient 4. Patient 4 is male and 68 years old at the date of first sample collection. HCC was clinically diagnosed after the date of final sample collection.

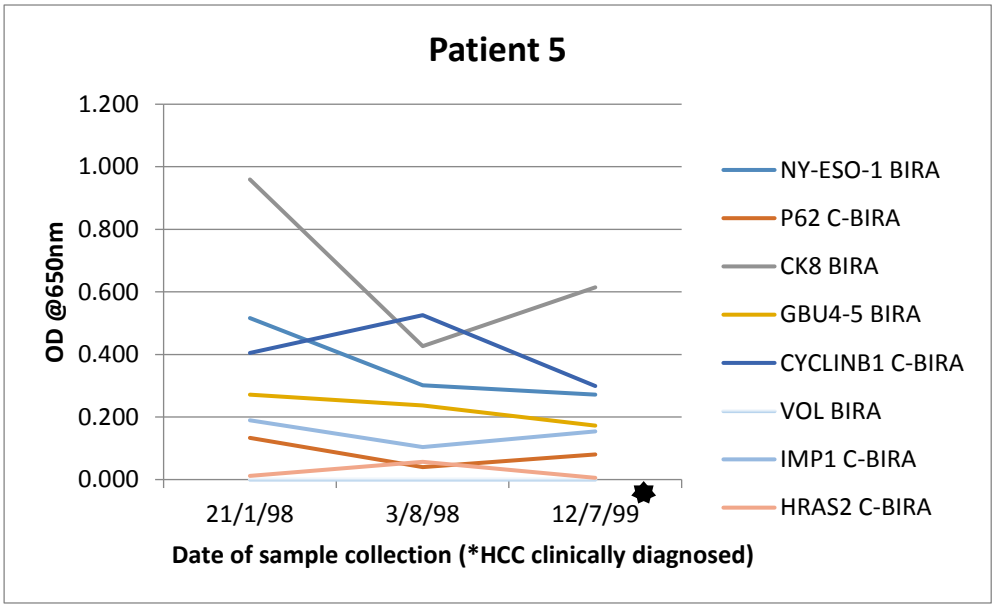


Figure 39 - OD readings at a protein concentration on 160nm of multiple samples from Patient 5. Patient 5 is female and 52 years old at the date of first sample collection. HCC was clinically diagnosed at the time the final sample was taken.

8.4 Discussion

This preliminary work was not intended to produce a statistically significant result or be the basis around which to form an algorithm, as that would

require a much larger sample set. Despite this, the results appear promising. There were AAb increases in 2-4 of the 5 patients leading up to HCC diagnosis. AAb rises in these patients occurred as early as 23 months before clinical HCC diagnosis, with rises in most patients occurring at least 6 months prior to clinical diagnosis. With such poor survival seen in anything past the earliest stage cancers, detection at these early time points could be key to increasing treatment success. There is also potential to apply this strategy to monitor cancer reoccurrence after treatment.

The hypothesis for this assay was to see low AAb levels in early time points and to see an increase in levels as the cancer develops, therefore reinforcing the idea that these AAb responses are cancer specific, this data is supportive of this hypothesis. This hypothesis warrants further testing and should be performed on the panel formed and validated in section Chapter 7. To establish a statistically significant result and creation of an algorithm to determine which increases show a high risk of cancer development, it will be necessary to monitor serial samples from patients at high risk of HCC to compare the AAb profiles of those patients who go on to develop HCC with those who do not. Unfortunately, it was not possible to collect the required samples within the time frame of this project.

Chapter 9. Discussion

This project has provided convincing evidence that AAbs are present in the serum of some patients with HCC. AAb levels are capable of separating HCC patients from high risk liver disease controls and healthy controls. The accuracy of a single AAb alone is limited, often with high specificity but low sensitivity. To overcome this, panels of AAbs are better than a single AAb alone at distinguishing HCC patients from controls. This is consistent with multiple reports in the literature[171,172,187].

In the training study, sensitivity reached 46% utilising a panel of 11 AAbs. The training study assessed 14 AAbs, so an increase in AAbs measured could increase sensitivity. However, sensitivity only reached 28% in the discovery assay when there were 42 AAbs measured. NRI analysis showed the marker positivity overlapped significantly between samples. This is highly likely as carcinogenesis usually requires multiple changes in cellular processes and therefore could produce more than one antigen capable of triggering an immune response.

The improvement between the discovery assay and the final training/validation panel came from two factors, antigen purity improvement and addition of markers which were not included in the discovery Assay. Even if all batches of antigen in the discovery assay had been produced to a much higher quality, it is unlikely that they would have increased the sensitivity significantly, as those that made it into the final panel, and reproduced in section Chapter 5, were able to show HCC/NCCLD differentiation even with HTP batches of antigen.

Specificity in the control groups was not 100%, an observation often seen with many blood-based tests for cancer. It is possible that some of this is due to contamination of cancers in the high-risk group (NCCLD). Average risk of HCC in high risk liver disease patients is in the region of 2-4% per year according to the AASLD guidelines[262]. Contamination in the NCCLD group at an assumed prevalence of 4% would lead to an actual specificity of 92.7% (specificity/(100-prevelance)) in the validation cohort. This result is of course only hypothetical

but may be clarified by follow-up of the patients from the NCCLD cohort. There is also a 6% difference between the specificities of the NCCLD controls and the healthy controls, 89% vs 95% respectively. As the risk in the general population is much lower (approximately 0.0048%), it is likely that the true specificity is closer to 89.2%, the 6% difference between the two control groups could be due to either undetected cancers or AAb presence related to liver disease.

Reactivity to MDM2 has previously been described by Liu et al. (2014)[193] to have the ability to distinguish HCC from liver disease controls with a sensitivity of 19.3% and specificity of 96.7%. That study reported 6.6% false positives in the chronic hepatitis control group and 0% false positives in the liver cirrhosis controls. In section Chapter 4 of this thesis, MDM2_C-BirA was run as part of the discovery assay. Results in section Chapter 4 showed a sensitivities of 1% at 97% specificity and 2% at 96% specificity. MDM2 produced in both studies was expressed using *E. coli* and purified using nickel affinity due to the 6xHIS tag. The observed differences could be due to sample set variation which again highlights the need for thorough validation of these tests. An alternative explanation could be due to the type of tag used. In this thesis all antigens contained a BirA tag which is an enzyme of bacterial origin. This tag may be bound by antibodies generated against bacterial infections in the patients examined, and there is evidence of increased anti-bacterial antibodies in liver disease patients[9,23,263,264]. Liu et al. used a Glutathione S-transferase (GST) tag which is derived from *Schistosoma japonicum*[265] and has been trialled as a vaccine for schistosomiasis[266] so is clearly capable of eliciting an immune response. Schistosomiasis is not endemic to Texas, the region Liu et al. obtained their samples from, so antibody levels against it should be low. One criticism of the Liu et al. study is that they did not control for levels of antibodies against the tag as was done with the VOL correction in this thesis. Whilst schistosomiasis is not endemic to Texas, it is in other regions of the world which have high rates of HCC, so this tag may not be broadly applicable. Along with the signal improvements seen in this thesis which are dependent on antigen batch purity, this highlights the issues with variation in AAb measurement

techniques. Standardisation of the methods could improve inter study variability.

An AAb based test could fit into the current clinical pathway in several ways. The first could be as a first line screening test. Currently Ultrasound with or without AFP measurement is the recommended screening test for high risk liver disease patients in the US[262]. The performance of ultrasound varies depending on centre and publication, however a meta-analysis shows the performance of ultrasound in high risk patients for screening to have 78% sensitivity and 89% specificity[267]. Sensitivity for early stage HCC however, is typically less than 60%[98]. AFP screening is often performed in addition to ultrasound to help detect some of the patients missed by ultrasound. The AAB+AFP test developed during this thesis could be beneficial as a replacement for AFP alone as it shows similar overall sensitivity but a higher sensitivity for early stage tumours than AFP. This improved detection of early stage tumours could be very complimentary to ultrasound which typically struggles to detect early stage tumours but can be very good at detecting larger late stage tumours. Unfortunately, this can only be established in a prospective trial or with a well-documented retrospective study that has follow-up on ultrasound negative patients. Such data was not available for the samples used in this thesis. The use of the test as a follow-up to ultrasound screening can however be modelled. The AASLD quotes the incidence of liver cancer as 2-4% in high risk patients in the US. Figure 40 shows the effect on the pathway if this test would be used to follow-up patients screened with ultrasound based on ultrasound performance of 78% sensitivity and 89% specificity[267] and a starting risk of HCC of 4%. Patients that test positive for ultrasound would have an increase in risk of HCC to 23%. If those ultrasound positive patients also had a positive result from the validation panel, the risk of HCC would further increase to 57.7% (2.5x increase). Any ultrasound negative patients with a positive AAb test would go from 1% risk to 4.5% (4.5x increase). This could be used to accelerate follow up of double positive patients and increase the screening frequency of ultrasound negative but AAb+AFP panel positive patients.

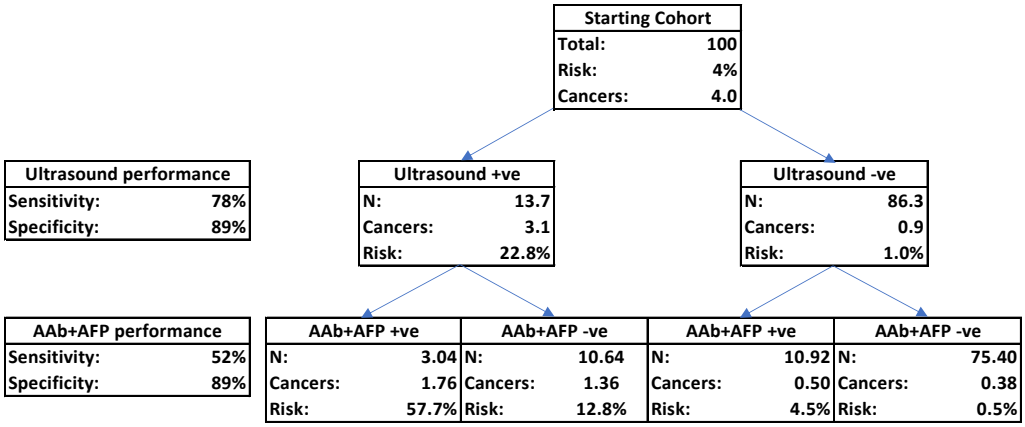


Figure 40 - Theoretical use of the AAb+AFP panel in a cohort of 100 high risk patients at a risk of HCC of 4%. Modelled increases in risk are based on the sensitivity and specificity of ultrasound and the AAb+AFP panel shown in the figure

Current Chinese guidelines recommend ultrasound plus AFP for HCC screening[268], however as China does not have a national healthcare system screening strategies vary by province. In rural areas AFP has been used as the primary screening tool[269] due to the cost and resource benefit over ultrasound. The performance of AFP is less than ideal, especially in early stage disease often showing less than <50% sensitivity[270]. Whilst the overall sensitivity of the AAb+AFP test is not high, specificity is high which lowers false positives and detection of BCLC stage A tumours is greater than 50%, a figure which AFP alone struggles to reach. As for BCLC stage 0, there was only one sample in the training and validation sets, so although negative, does not give a significant representation of the performance of this test in BLCL 0 patients. This test could therefore be used in place of AFP in regions where ultrasound screening is prohibitive.

In the US, the AASLD guidelines recommended further investigation of ultrasound detected liver nodules with contrast enhanced MRI or multiphasic CT. The guidelines recommend avoiding biopsy where possible if a diagnosis can be made by imax`ging. To do so they recommend that imaging indicative of HCC from both CT and MRI can be used as to diagnose HCC and forgo biopsy. This can become expensive and results in a drop of sensitivity to 33%. The AAb+AFP test could be used as a rule-in test for follow-up treatment rather

than using secondary imaging. When modelled on the sensitivity and specificity for CT and MRI quoted in the guidelines as well as the sensitivity and specificity for ultrasound screening quoted by Chou et al.[271], the risk pathway would be as seen in Figure 41. A positive MRI or CT result would put a patient at a risk of 81% and 84% respectively. If that patient then went on to have a positive result from the AAb+AFP test, risk would increase to 95% and 96% respectively. This risk may be significantly high enough to avoid the need for biopsy. Any AAB+AFP negative tumours could go on to have further imaging, this could help reduce the number of expensive scans needed. The test could also be used to identify risk in any patients that receive a negative result from CT or MRI scanning, these patients could then be sent down an accelerated pathway for further imaging.

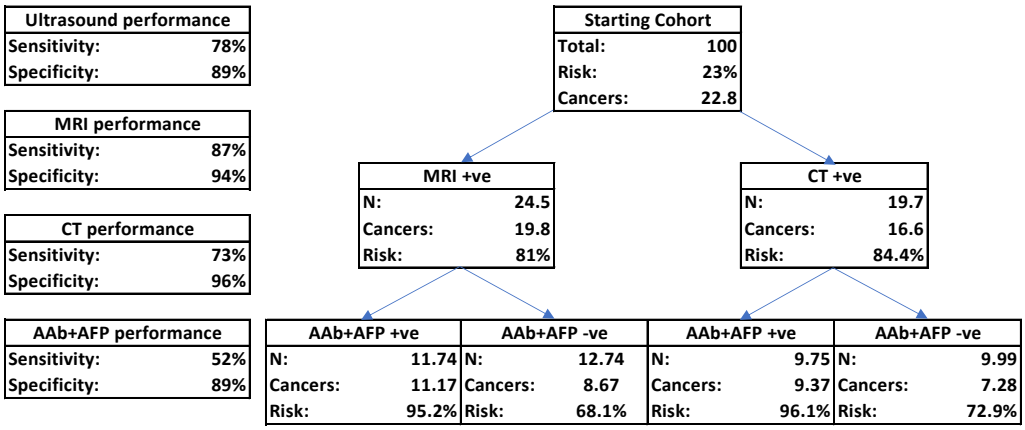


Figure 41 - Theoretical use of the AAb+AFP panel in a cohort of 100 patients with a liver lesion identified by Ultrasound screening at a HCC risk of 23%. Modelled increases in risk are based on the sensitivity and specificity of ultrasound, CT, MRI and the AAb+AFP panel shown in the figure

Whilst there are many options for how this panel could be utilised, its performance is still far from perfect. At the current sensitivity of 52% this test would currently miss half of cancers so therefore may not be suitable to be the only method of screening, however as the specificity is high it may be best suited to a rule in test where any negative results undergo traditional follow-up rather than being ruled out. On the contrary, whilst specificity is high at 89%, if this test were to be used for screening, this may not be high enough. As HCC has a relatively low prevalence, even an 11% false positive rate is enough to

significantly effect PPV as there are more chances to give false positive results than true positive results. At the 4% prevalence modelled earlier, PPV would only be 16.5% and therefore an 83.5% chance that a positive result is a false positive. However, in a post US screened population, where prevalence is higher, (23% as modelled earlier), PPV would increase to 58.5%. This stresses the need for correct application of the test, or improvement of the test's overall performance. At a single cut-off, the AUC of a test can be calculated as $\frac{sensitivity+specificity}{2}$, giving the AAb+AFP panel an AUC of 0.705. AUCs are commonly ranked as; <0.5 is not useful, 0.5-0.6 is bad, 0.6-0.7 is sufficient, 0.7-0.8 is good, 0.8-0.9 is very good and 0.9-1.0 is excellent[272]. This current test would be classified as good but has a lot of room for improvement. There are three primary ways in which this test could be improve:

(i) Whilst not necessarily improving the overall accuracy of the test, it may be possible to optimise the situation in which it is used in to achieve more desired characteristics. By increasing the population prevalence via stratification, it is possible to increase the PPV with a trade off in lower NPV. There are many risk stratification tests for liver diseases which attempt to predict the chance of HCC[273]. These are often combinations of commonly measured patient factors such as the PAGE-B score which determines 5-year HCC risk in HBV patients based on age, gender and platelet count[274]. Use of these to stratify high risk patients before using the AAb+AFP test, could increase the prevalence of the disease in the testing population by ruling out patients at lower risk of HCC. An increased PPV will benefit if the test is to be used in situations where follow-up may have a chance of harm.

(ii) Serial sampling of patients to create individual profiles rather than a population-based cut-off may offer a chance to improve both sensitivity and specificity. Chapter 8 showed some basic evidence that this may be possible however, requires much further research. This could be implemented in a similar style to the ROCA algorithm for CA-125 screening of ovarian cancer[175]. Analysis of patient specific profiles may reduce false positives from patients without cancer who enter screening with naturally high levels of some AAbs

and increase sensitivity by detecting minor rises in patients with HCC that would not normally cross a single population-based cut-off. The trade-off for this, is the extensive validation required plus the increased cost of regular repeat screening for all patients which may make the costs prohibitive in health economic studies.

(iii) Addition of new markers may be the best way to increase the performance of the panel. However, the answer to this may not be to add additional AAbs to the panel, but markers of other biological processes. Multiple biological pathways are affected during the development of cancer and many of these leave behind detectable markers. This thesis has already looked at combining AAbs with a circulating protein biomarker (AFP). There may be more circulating protein biomarkers such as DCP, GPC3 and GP73 which may further increase the sensitivity of the panel[275]. In addition to AAbs and circulating proteins, it is also possible to measure changed expression levels of metabolites, miRNAs and lipids as well as DNA mutations and methylation changes from circulating tumour DNA and circulating tumour cells. Cohen et al. showed that detection of KRAS mutations in circulating tumour DNA could increase the sensitivity of a 4 protein panel for pancreatic cancer from 54% to 64% with a minor (0.5%) reduction in specificity[276]. Whilst addition of these markers may increase the sensitivity of a test, this will often be at the trade-off of reduced specificity as many biomarkers are not 100% specific. This requires careful cut-off selection for each marker in the panel or formation of an algorithm. As many of the markers are analysed via different techniques, the more that are added to the panel increases both the complexity and cost of the test and therefore lowering the chance it will be cost effective. This will also require higher quality of sample processing, storage and handling so that the samples are suitable for use in multiple techniques, this is especially critical in miRNA studies[277]

One of the oldest circulating tumour markers to measure is cancer related proteins and is the reason that most of the current validated cancer biomarkers are circulating proteins, e.g. AFP for HCC. These are usually discovered using 2D gel electrophoresis followed by mass spectrometry analysis and are then

primarily measured via sandwich ELISA. Pre-designed microarray and bead-based technologies are becoming the new standard for marker discovery as these can probe approximately 50 proteins at once and require little downstream work to create an assay but require specialist lab equipment which creates potential barriers for deployment in hospital labs[278,279]. Metabolomics are one of the newest areas of biomarker research, especially for diseases of organs with important metabolic functions. Volatile organic compounds present in breath have become a key area of research for the detection of lung cancer[280]. The liver is a key organ in mediating the levels of numerous metabolites, making the study of them a potential biomarker of HCC[281]. Accessible metabolomic biomarkers for screening purposes can be detected in the blood, urine and faeces and are primarily detected via mass spectrometry. Luo et al. found a 2 metabolite panel of phenylalanyl-tryptophan and glycocholate to have a sensitivity of 92% and specificity of 53% for distinguishing HCC from cirrhotic controls[282]. Circulating tumour DNA, miRNA, and DNA methylation dysregulation can be detected in the serum of many cancer patients. New biomarkers tend to be discovered via microarray analysis of tumour tissue with microarrays containing thousands of probes[283]. Any leads are then converted for measurement using quantitative PCR due to the prohibitive costs of the microarrays. Whilst these elements are often present at low concentrations in blood, it is possible to amplify the signals and therefore making detection easier but at the trade-off of introducing another variable into the assay. These tests have potential to give more information about the underlying tumour which may correlate with treatment strategies and prognosis, however presence may be correlated with later tumour stage[284]. Combination with any of the above biomarker types maybe be beneficial to the current AAb+AFP panel if they can increase the accuracy of the panel at a reasonable cost effectiveness, cost increase is not necessarily prohibitive if the gains can out way the costs in health economic studies[284].

This project has established a panel of AAbs and AFP which could be used at multiple stages of the clinical pathway. However, to decide just how effective

this could be will require multiple further studies comparing this against current clinical pathways. The patients from the training and validation set should be followed up to establish the true rates of cancer in that group and determine the actual specificity. Evidence from this thesis and from the literature shows variation in how important it is to thoroughly validate these tests. This needs to be shown in multiple areas. The usefulness should be shown in independent sample sets to increase the confidence that the results are genuine and not sample set specific. These samples should ideally come from studies that represent the proposed clinical use. For example, most studies, including those in this thesis, are based on case/control sample sets obtained retrospectively. If the intended use is screening, the samples should come from a prospective collection of samples from the intended screening group, in this case, patients with chronic liver disease. This will show the true utility of such tests. Extensive work must also be performed to validate the repeatability and stability of the test, to ensure sufficient confidence in the results.

Whilst this thesis has not used prospective high-risk sample collections or shown thorough validation of the assay method, it does accomplish an important step in test development. Studies such as this are important as a test must be determined before proceeding with the further validation steps. This thesis shows sufficient evidence that these studies should be performed and will be the next course of action for this work. There is also potential to develop a serial measurement algorithm, however this will require a well-planned study with a large sample set.

Chapter 10. References

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* (2018). doi:10.3322/caac.21492
2. Molina, D. K. & DiMaio, V. J. M. Normal organ weights in men: part II-the brain, lungs, liver, spleen, and kidneys. *Am. J. Forensic Med. Pathol.* **33,** 368–72 (2012).
3. Molina, D. K. & DiMaio, V. J. M. Normal Organ Weights in Women: Part II-The Brain, Lungs, Liver, Spleen, and Kidneys. *Am. J. Forensic Med. Pathol.* **36,** 182–7 (2015).
4. Gulec, S. A., Sztejnberg, M. L., Siegel, J. A., Jevremovic, T. & Stabin, M. Hepatic structural dosimetry in (90)Y microsphere treatment: a Monte Carlo modeling approach based on lobular microanatomy. *J. Nucl. Med.* **51,** 301–10 (2010).
5. Wu, G. Amino acids: metabolism, functions, and nutrition. *Amino Acids* **37,** 1–17 (2009).
6. Iyer, K. R. & Sinz, M. W. Characterization of Phase I and Phase II hepatic drug metabolism activities in a panel of human liver preparations. *Chem. Biol. Interact.* **118,** 151–169 (1999).
7. Kamdem, L. K., Meineke, I., Gödtel-Armbrust, U., Brockmöller, J. & Wojnowski, L. Dominant contribution of P450 3A4 to the hepatic carcinogenic activation of aflatoxin B1. *Chem. Res. Toxicol.* **19,** 577–86 (2006).
8. Boon, A.-C., Bulmer, A. C., Coombes, J. S. & Fassett, R. G. Circulating bilirubin and defense against kidney disease and cardiovascular mortality: mechanisms contributing to protection in clinical investigations. *Am. J. Physiol. Renal Physiol.* **307,** F123-36 (2014).
9. Jeong, W.-I. & Gao, B. Innate immunity and alcoholic liver fibrosis. *J. Gastroenterol. Hepatol.* **23 Suppl 1,** S112-8 (2008).
10. Livingstone, C. Insulin-like growth factor-I (IGF-I) and clinical nutrition. *Clin. Sci. (Lond).* **125,** 265–80 (2013).
11. Hitchcock, I. S. & Kaushansky, K. Thrombopoietin from beginning to end. *Br. J. Haematol.* **165,** 259–68 (2014).
12. Lubel, J. S., Herath, C. B., Burrell, L. M. & Angus, P. W. Liver disease and the renin-angiotensin system: recent discoveries and clinical implications. *J. Gastroenterol. Hepatol.* **23,** 1327–38 (2008).
13. Postic, C., Dentin, R. & Girard, J. Role of the liver in the control of carbohydrate and lipid homeostasis. *Diabetes Metab.* **30,** 398–408 (2004).
14. Rinella, M. E. Nonalcoholic Fatty Liver Disease. *JAMA* **313,** 2263 (2015).

15. Michalopoulos, G. K. Liver regeneration. *J. Cell. Physiol.* **213,** 286–300 (2007).

16. Jelic, S. & Sotiropoulos, G. C. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **21 Suppl 5,** v59-64 (2010).

17. Gao, B., Jeong, W.-I. & Tian, Z. Liver: An organ with predominant innate immunity. *Hepatology* **47,** 729–36 (2008).

18. Roncalli, M., Park, Y. N. & Di Tommaso, L. Histopathological classification of hepatocellular carcinoma. *Dig. Liver Dis.* **42,** S228–S234 (2010).

19. Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur. J. Cancer* **49,** 1374–403 (2013).

20. Hartwell, H. J., Petrosky, K. Y., Fox, J. G., Horseman, N. D. & Rogers, A. B. Prolactin prevents hepatocellular carcinoma by restricting innate immune activation of c-Myc in mice. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 11455–60 (2014).

21. Fattovich, G., Stroffolini, T., Zagni, I. & Donato, F. Hepatocellular carcinoma in cirrhosis: Incidence and risk factors. *Gastroenterology* **127,** S35–S50 (2004).

22. Ramakrishna, G. *et al.* From cirrhosis to hepatocellular carcinoma: new molecular insights on inflammation and cellular senescence. *Liver cancer* **2,** 367–83 (2013).

23. Møller, S., Henriksen, J. H. & Bendtsen, F. Extrahepatic complications to cirrhosis and portal hypertension: haemodynamic and homeostatic aspects. *World J. Gastroenterol.* **20,** 15499–517 (2014).

24. Kim, M. Y. *et al.* Hepatic venous pressure gradient can predict the development of hepatocellular carcinoma and hyponatremia in decompensated alcoholic cirrhosis. *Eur. J. Gastroenterol. Hepatol.* **21,** 1241–6 (2009).

25. Garcia-Tsao, G., Parikh, C. R. & Viola, A. Acute kidney injury in cirrhosis. *Hepatology* **48,** 2064–77 (2008).

26. El-Serag, H. B. & Rudolph, K. L. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* **132,** 2557–76 (2007).

27. Wong, R. J. Obesity and non-alcoholic fatty liver disease: Disparate associations among Asian populations. *World J. Hepatol.* **6,** 263 (2014).

28. Williams, K. H., Shackel, N. A., Gorrell, M. D., McLennan, S. V. & Twigg, S. M. Diabetes and Nonalcoholic Fatty Liver Disease: A Pathogenic Duo. *Endocr. Rev.* **34,** 84–129 (2013).

29. Fon Tacer, K. & Rozman, D. Nonalcoholic Fatty liver disease: focus on lipoprotein and lipid deregulation. *J. Lipids* **2011,** 783976 (2011).

30. Cameron, A. J., Shaw, J. E. & Zimmet, P. Z. The metabolic syndrome: prevalence in worldwide populations. *Endocrinol. Metab. Clin. North Am.* **33,** 351–75, table of contents (2004).

31. Landskron, G., De la Fuente, M., Thuwajit, P., Thuwajit, C. & Hermoso, M. A. Chronic inflammation and cytokines in the tumor microenvironment. *J. Immunol. Res.* **2014,** 149185 (2014).

32. Drutskaya, M. S., Efimov, G. A., Kruglov, A. A., Kuprash, D. V & Nedospasov, S. A. Tumor necrosis factor, lymphotoxin and cancer. *IUBMB Life* **62,** 283–9 (2010).

33. Dowman, J. K., Tomlinson, J. W. & Newsome, P. N. Systematic review: the diagnosis and staging of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. *Aliment. Pharmacol. Ther.* **33,** 525–40 (2011).

34. Purohit, V., Gao, B. & Song, B.-J. Molecular mechanisms of alcoholic fatty liver. *Alcohol. Clin. Exp. Res.* **33,** 191–205 (2009).

35. Matsuda, T. *et al.* Effective utilization of N2-ethyl-2'-deoxyguanosine triphosphate during DNA synthesis catalyzed by mammalian replicative DNA polymerases. *Biochemistry* **38,** 929–35 (1999).

36. Brooks, P. J. & Theruvathu, J. A. DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* **35,** 187–93 (2005).

37. Dupont, I., Lucas, D., Clot, P., Ménez, C. & Albano, E. Cytochrome P4502E1 inducibility and hydroxyethyl radical formation among alcoholics. *J. Hepatol.* **28,** 564–571 (1998).

38. Svegliati-Baroni, G. *et al.* Early response of alpha2(I) collagen to acetaldehyde in human hepatic stellate cells is TGF-beta independent. *Hepatology* **42,** 343–52 (2005).

39. Wu, F., Wang, J., Pu, C., Qiao, L. & Jiang, C. Wilson's disease: a comprehensive review of the molecular mechanisms. *Int. J. Mol. Sci.* **16,** 6419–31 (2015).

40. Kowdley, K. V. Iron, hemochromatosis, and hepatocellular carcinoma. *Gastroenterology* **127,** S79–S86 (2004).

41. Wong, R. J., Gish, R., Frederick, T., Bzowej, N. & Frenette, C. Development of hepatocellular carcinoma in autoimmune hepatitis patients: a case series. *Dig. Dis. Sci.* **56,** 578–85 (2011).

42. El-Serag, H. B. & Rudolph, K. L. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* **132,** 2557–76 (2007).

43. Yim, H. J. & Lok, A. S.-F. Natural history of chronic hepatitis B virus infection: what we knew in 1981 and what we know in 2005. *Hepatology* **43,** S173-81 (2006).

44. Trépo, C., Chan, H. L. Y. & Lok, A. Hepatitis B virus infection. *Lancet* **384,** 2053–63 (2014).

45. Kramvis, A. Genotypes and genetic variability of hepatitis B virus. *Intervirology* **57,** 141–50 (2014).

46. Kondo, Y., Ninomiya, M., Kakazu, E., Kimura, O. & Shimosegawa, T. Hepatitis B surface antigen could contribute to the immunopathogenesis of hepatitis B virus infection. *ISRN Gastroenterol.* **2013,** 935295 (2013).

47. Feitelson, M. A. & Duan, L. X. Hepatitis B virus X antigen in the pathogenesis of chronic infections and the development of hepatocellular carcinoma. *Am. J. Pathol.* **150,** 1141–57 (1997).

48. Sun, C.-A. Incidence and Cofactors of Hepatitis C Virus-related Hepatocellular Carcinoma: A Prospective Study of 12,008 Men in Taiwan. *Am. J. Epidemiol.* **157,** 674–682 (2003).

49. Mohd Hanafiah, K., Groeger, J., Flaxman, A. D. & Wiersma, S. T. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* **57,** 1333–42 (2013).

50. Gupta, S. Test Characteristics of α-Fetoprotein for Detecting Hepatocellular Carcinoma in Patients with Hepatitis C. *Ann. Intern. Med.* **139,** 46 (2003).

51. Seeff, L. B. Natural history of chronic hepatitis C. *Hepatology* **36,** s35–s46 (2002).

52. Barth, H. Hepatitis C virus: Is it time to say goodbye yet? Perspectives and challenges for the next decade. *World J. Hepatol.* **7,** 725–37 (2015).

53. Roberta, D., Cristina, D. & Massimo, C. Hepatocellular Carcinoma in Patients with a Sustained Response to Anti-Hepatitis C Therapy. *Int. J. Mol. Sci.* **16,** 19698–19712 (2015).

54. Jeong, S. W., Jang, J. Y. & Chung, R. T. Hepatitis C virus and hepatocarcinogenesis. *Clin. Mol. Hepatol.* **18,** 347 (2012).

55. Tsutsumi, T. Hepatitis C virus core protein activates ERK and p38 MAPK in cooperation with ethanol in transgenic mice. *Hepatology* **38,** 820–828 (2003).

56. Holder, K. A., Russell, R. S. & Grant, M. D. Natural Killer Cell Function and Dysfunction in Hepatitis C Virus Infection. *Biomed Res. Int.* **2014,** 1–9 (2014).

57. Kwun, H. J., Jang, K. L., Lee, M. N., Jung, E. Y. & Ahn, J. Y. p53-dependent transcriptional repression of p21waf1 by hepatitis C virus NS3. *J. Gen. Virol.* **82,** 2235–2241 (2001).

58. Chung, Y.-L., Sheu, M.-L. & Yen, S.-H. Hepatitis C virus NS5A as a potential viral Bcl-2 homologue interacts with Bax and inhibits apoptosis in hepatocellular carcinoma. *Int. J. Cancer* **107,** 65–73 (2003).

59. Jung, S.-Y. & Kim, Y.-J. C-terminal region of HBx is crucial for mitochondrial DNA damage. *Cancer Lett.* **331,** 76–83 (2013).

60. Zhu, Y.-Z. *et al.* Hepatitis B virusXprotein promotes hypermethylation of p16INK4A promoter through upregulation of DNA methyltransferases in hepatocarcinogenesis. *Exp. Mol. Pathol.* **89,** 268–275 (2010).

61. Ayub, A., Ashfaq, U. A. & Haque, A. HBV induced HCC: major risk factors from genetic to molecular level. *Biomed Res. Int.* **2013,** 810461 (2013).

62. Stern, M. C. *et al.* Hepatitis B, Aflatoxin B1, and p53 Codon 249 Mutation in Hepatocellular Carcinomas from Guangxi, People's Republic of China, and a Meta-analysis of Existing Studies. *Cancer Epidemiol. Biomarkers Prev.* **10,** 617–625 (2001).

63. Yang, W.-S. *et al.* The Role of Pre-Existing Diabetes Mellitus on Hepatocellular Carcinoma Occurrence and Prognosis: A Meta-Analysis of Prospective Cohort Studies. *PLoS One* **6,** e27326 (2011).

64. Eliyatkın, N., Yalçın, E., Zengel, B., Aktaş, S. & Vardar, E. Molecular Classification of Breast Carcinoma: From Traditional, Old-Fashioned Way to A New Age, and A New Way. *J. breast Heal.* **11,** 59–66 (2015).

65. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5,** 2929–43 (2015).

66. Slamon, D. *et al.* Adjuvant Trastuzumab in HER2-Positive Breast Cancer. *N. Engl. J. Med.* **365,** 1273–1283 (2011).

67. Nault, J.-C. & Zucman-Rossi, J. TERT promoter mutations in primary liver tumors. *Clin. Res. Hepatol. Gastroenterol.* **40,** 9–14 (2016).

68. Nault, J. C. *et al.* Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology* **60,** 1983–1992 (2014).

69. Kunst, C., Haderer, M., Heckel, S., Schlosser, S. & Müller, M. The p53 family in hepatocellular carcinoma. *Transl. Cancer Res.* **5,** 632–638 (2016).

70. Olivier, M. *et al.* The IARC TP53 database: New online mutation analysis and recommendations to users. *Hum. Mutat.* **19,** 607–614 (2002).

71. Bojesen, A., Hertz, J. M. & Gravholt, C. H. Genotype and phenotype in Klinefelter syndrome - impact of androgen receptor polymorphism and skewed X inactivation.

72. Gouas, D. & Shi, H. The aflatoxin-induced TP53 mutation at codon 249 (R249S): Biomarker of exposure, early detection and target for therapy. *Cancer Lett.* **286,** 29–37 (2009).

73. Thompson, M. D. & Monga, S. P. S. WNT/β-catenin signaling in liver health and disease. *Hepatology* **45,** 1298–1305 (2007).

74. Tornesello, M. L. *et al.* Mutations in TP53, CTNNB1 and PIK3CA genes in hepatocellular carcinoma associated with hepatitis B and hepatitis C virus infections. *Genomics* **102,** 74–83 (2013).

75. Cieply, B., Zeng, G., Proverbs-Singh, T., Geller, D. A. & Monga, S. P. S. Unique phenotype of hepatocellular cancers with exon-3 mutations in beta-catenin gene. *Hepatology* **49,** 821–831 (2009).

76. Hsieh, A., Kim, H.-S., Lim, S.-O., Yu, D.-Y. & Jung, G. Hepatitis B viral X protein interacts with tumor suppressor adenomatous polyposis coli to activate Wnt/β-catenin signaling. *Cancer Lett.* **300,** 162–172 (2011).

77. Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44,** 694–698 (2012).

78. Wang, K. *et al.* Genomic landscape of copy number aberrations enables the identification of oncogenic drivers in hepatocellular carcinoma. *Hepatology* **58,** 706–717 (2013).

79. Lee, J. W. *et al.* PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. *Oncogene* **24,** 1477–1480 (2005).

80. Colombino, M. *et al.* BRAF and PIK3CA genes are somatically mutated in hepatocellular carcinoma among patients from South Italy. *Cell Death Dis.* **3,** e259–e259 (2012).

81. Tanaka, Y. *et al.* Absence of PIK3CA hotspot mutations in hepatocellular carcinoma in Japanese patients. *Oncogene* **25,** 2950–2952 (2006).

82. Nault, J.-C., Galle, P. R. & Marquardt, J. U. The role of molecular enrichment on future therapies in hepatocellular carcinoma. *J. Hepatol.* **69,** 237–247 (2018).

83. Llovet, J. M. *et al.* Sorafenib in Advanced Hepatocellular Carcinoma. *N. Engl. J. Med.* **359,** 378–390 (2008).

84. Heo, Y.-A. & Syed, Y. Y. Regorafenib: A Review in Hepatocellular Carcinoma. *Drugs* **78,** 951–958 (2018).

85. Rimassa, L. *et al.* Tivantinib for second-line treatment of MET-high, advanced hepatocellular carcinoma (METIV-HCC): a final analysis of a phase 3, randomised, placebo-controlled study. *Lancet. Oncol.* **19,** 682–693 (2018).

86. Bruix, J. & Sherman, M. Management of hepatocellular carcinoma: an update. *Hepatology* **53,** 1020–2 (2011).

87. Song, P. *et al.* Biomarkers: evaluation of screening for and early diagnosis of hepatocellular carcinoma in Japan and china. *Liver cancer* **2,** 31–9 (2013).

88. Zhang, B.-H., Yang, B.-H. & Tang, Z.-Y. Randomized controlled trial of screening for hepatocellular carcinoma. *J. Cancer Res. Clin. Oncol.* **130,** 417–22 (2004).

89. Tong, M. J., Sun, H.-E., Hsien, C. & Lu, D. S. K. Surveillance for Hepatocellular Carcinoma Improves Survival in Asian-American Patients with Hepatitis B: Results from a Community-Based Clinic. *Dig. Dis. Sci.* **55,** 826–835 (2010).

90. Huo, T.-I. *et al.* Survival impact of delayed treatment in patients with hepatocellular carcinoma undergoing locoregional therapy: Is there a lead-time bias? *Scand. J. Gastroenterol.* **42,** 485–492 (2007).

91. Ruggeri, M. Hepatocellular carcinoma: cost-effectiveness of screening. A systematic review. *Risk Manag. Healthc. Policy* **5,** 49–54 (2012).

92. Hu, B., Tian, X., Sun, J. & Meng, X. Evaluation of Individual and Combined Applications of Serum Biomarkers for Diagnosis of Hepatocellular Carcinoma: A Meta-Analysis. *Int. J. Mol. Sci.* **14,** 23559–23580 (2013).

93. Ge, T. *et al.* Diagnostic values of alpha-fetoprotein, dickkopf-1, and osteopontin for hepatocellular carcinoma. *Med. Oncol.* **32,** 59 (2015).

94. Jia, X., Liu, J., Gao, Y., Huang, Y. & Du, Z. Diagnosis Accuracy of Serum Glypican-3 in Patients with Hepatocellular Carcinoma: A Systematic Review with Meta-analysis. *Arch. Med. Res.* **45,** 580–588 (2014).

95. Berhane, S. *et al.* Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. (2016). doi:10.1016/j.cgh.2015.12.042

96. Tzartzeva, K. *et al.* Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. (2018). doi:10.1053/j.gastro.2018.01.064

97. El-Serag, H. B. & Davila, J. A. Surveillance for hepatocellular carcinoma: in whom and how? *Therap. Adv. Gastroenterol.* **4,** 5–10 (2011).

98. Singal, A. *et al.* Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. *Aliment. Pharmacol. Ther.* **30,** 37–47 (2009).

99. TATARINOV, I. S. [DETECTION OF EMBRYO-SPECIFIC ALPHA-GLOBULIN IN THE BLOOD SERUM OF A PATIENT WITH PRIMARY LIVER CANCER]. *Vopr. meditsinskoĭ khimii* **10,** 90–1

100. Hu, B., Tian, X., Sun, J. & Meng, X. Evaluation of individual and combined applications of serum biomarkers for diagnosis of hepatocellular carcinoma: a meta-analysis. *Int. J. Mol. Sci.* **14,** 23559–80 (2013).

101. Farinati, F. *et al.* Diagnostic and prognostic role of alpha-fetoprotein in hepatocellular carcinoma: both or neither? *Am. J. Gastroenterol.* **101,** 524–32 (2006).

102. Liu, C. *et al.* Value of α-fetoprotein in association with clinicopathological features of hepatocellular carcinoma. *World J. Gastroenterol.* **19,** 1811–9 (2013).

103. Li, D., Mallory, T. & Satomura, S. AFP-L3: a new generation of tumor marker for hepatocellular carcinoma. *Clin. Chim. Acta* **313,** 15–19 (2001).

104. Yi, X., Yu, S. & Bao, Y. Alpha-fetoprotein-L3 in hepatocellular carcinoma: a meta-analysis. *Clin. Chim. Acta.* **425,** 212–20 (2013).

105. Suzuki, M. *et al.* Des-gamma-carboxy prothrombin is a potential autologous growth factor for hepatocellular carcinoma. *J. Biol. Chem.* **280,** 6409–15 (2005).

106. Zhu, R. *et al.* Diagnostic Performance of Des-γ-carboxy Prothrombin for Hepatocellular Carcinoma: A Meta-Analysis. *Gastroenterol. Res. Pract.* **2014,** 529314 (2014).

107. Zhou, Y., Yin, X., Ying, J. & Zhang, B. Golgi protein 73 versus alpha-fetoprotein as a biomarker for hepatocellular carcinoma: a diagnostic meta-analysis. *BMC Cancer* **12,** 17 (2012).

108. Cabiati, M. *et al.* Osteopontin in hepatocellular carcinoma: A possible biomarker for diagnosis and follow-up. *Cytokine* **99,** 59–65 (2017).

109. Li, G. *et al.* Identification of circulating MicroRNAs as novel potential biomarkers for hepatocellular carcinoma detection: a systematic review and meta-analysis. *Clin. Transl. Oncol.* (2015). doi:10.1007/s12094-015-1294-y

110. Wang, W. *et al.* MicroRNA-21 and the clinical outcomes of various

carcinomas: a systematic review and meta-analysis. *BMC Cancer* **14,** 819 (2014).

111. Witwer, K. W. Data submission and quality in microarray-based microRNA profiling. *Clin. Chem.* **59,** 392–400 (2013).

112. MacLellan, S. A., MacAulay, C., Lam, S. & Garnis, C. Pre-profiling factors influencing serum microRNA levels. *BMC Clin. Pathol.* **14,** 27 (2014).

113. Berhane, S. *et al.* Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. *Clin. Gastroenterol. Hepatol.* **14,** 875–886.e6 (2016).

114. Toyoda, H. *et al.* Clinical Practice Guidelines for Hepatocellular Carcinoma Differ between Japan, United States, and Europe. *Liver cancer* **4,** 85–95 (2015).

115. Yu, N. C. *et al.* CT and MRI improve detection of hepatocellular carcinoma, compared with ultrasound alone, in patients with cirrhosis. *Clin. Gastroenterol. Hepatol.* **9,** 161–7 (2011).

116. Oliver, J. H., Baron, R. L., Federle, M. P. & Rockette, H. E. Detecting hepatocellular carcinoma: value of unenhanced or arterial phase CT imaging or both used in conjunction with conventional portal venous phase contrast-enhanced CT imaging. *AJR. Am. J. Roentgenol.* **167,** 71–7 (1996).

117. Bialecki, E. S. & Di Bisceglie, A. M. Diagnosis of hepatocellular carcinoma. *HPB (Oxford).* **7,** 26–34 (2005).

118. Digumarthy, S. R., Sahani, D. V & Saini, S. MRI in detection of hepatocellular carcinoma (HCC). *Cancer Imaging* **5,** 20–4 (2005).

119. Willatt, J., Ruma, J. A., Azar, S. F., Dasika, N. L. & Syed, F. Imaging of hepatocellular carcinoma and image guided therapies - how we do it. *Cancer Imaging* **17,** 9 (2017).

120. Borzio, M. *et al.* The evaluation of fine-needle procedures for the diagnosis of focal liver lesions in cirrhosis. *J. Hepatol.* **20,** 117–121 (1994).

121. Pons, F., Varela, M. & Llovet, J. M. Staging systems in hepatocellular carcinoma. *HPB (Oxford).* **7,** 35–41 (2005).

122. Cillo, U. *et al.* The critical issue of hepatocellular carcinoma prognostic classification: which is the best tool available? *J. Hepatol.* **40,** 124–131 (2004).

123. Guglielmi, A. *et al.* Hepatocellular carcinoma: surgical perspectives beyond the barcelona clinic liver cancer recommendations. *World J. Gastroenterol.* **20,** 7525–33 (2014).

124. Mazzaferro, V. *et al.* Milan criteria in liver transplantation for hepatocellular carcinoma: an evidence-based analysis of 15 years of experience. *Liver Transpl.* **17 Suppl 2,** S44-57 (2011).

125. Piscaglia, F. *et al.* A new priority policy for patients with hepatocellular carcinoma awaiting liver transplantation within the model for end-stage liver disease system. *Liver Transpl.* **13,** 857–66 (2007).

126. Wald, C. *et al.* New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. *Radiology* **266,** 376–82 (2013).

127. Wáng, Y.-X. J., De Baere, T., Idée, J.-M. & Ballet, S. Transcatheter embolization therapy in liver cancer: an update of clinical evidences. *Chinese Journal of Cancer Research* **27,** 96–121

128. Bruix, J. *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet (London, England)* **389,** 56–66 (2017).

129. Kurosaki, T., Kometani, K. & Ise, W. Memory B cells. *Nat. Rev. Immunol.* **15,** 149–159 (2015).

130. Forthal, D. N. Functions of Antibodies. *Microbiol. Spectr.* **2,** 1–17 (2014).

131. Vyas, J. M., Van der Veen, A. G. & Ploegh, H. L. The known unknowns of antigen processing and presentation. *Nat. Rev. Immunol.* **8,** 607–618 (2008).

132. Kloetzel, P.-M. The proteasome and MHC class I antigen processing. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695,** 225–233 (2004).

133. Jensen, P. E. Recent advances in antigen processing and presentation. *Nat. Immunol.* **8,** 1041–1048 (2007).

134. Wieczorek, M. *et al.* Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front. Immunol.* **8,** 292 (2017).

135. Roche, P. A. & Cresswell, P. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature* **345,** 615–618 (1990).

136. Zhang, Y., Garcia-Ibanez, L. & Toellner, K.-M. Regulation of germinal center B-cell differentiation. *Immunol. Rev.* **270,** 8–19 (2016).

137. Shipman, W. D., Dasoveanu, D. C. & Lu, T. T. Tertiary lymphoid organs in systemic autoimmune diseases: pathogenic or protective? *F1000Research* **6,** 196 (2017).

138. Harwood, N. E. & Batista, F. D. Antigen presentation to B cells. *F1000 Biol. Rep.* **2,** 87 (2010).

139. Nagata, S. & Pastan, I. Removal of B cell epitopes as a practical approach for reducing the immunogenicity of foreign protein-based therapeutics. *Adv. Drug Deliv. Rev.* **61,** 977–85 (2009).

140. Ovary, Z. & Benacerraf, B. Immunological Specificity of the Secondary Response with Dinitrophenylated Proteins. *Exp. Biol. Med.* **114,** 72–76 (1963).

141. Greenbaum, J. A. *et al.* Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J. Mol. Recognit.* **20,** 75–82 (2007).

142. Regenmortel, M. H. V. What Is a B-Cell Epitope? in 3–20 (Humana Press, 2009). doi:10.1007/978-1-59745-450-6_1

143. Shlomchik, M. J. Sites and stages of autoreactive B cell activation and regulation. *Immunity* **28,** 18–28 (2008).

144. Fishman, J. M., Wiles, K. & Wood, K. J. The Acquired Immune System Response to Biomaterials, Including Both Naturally Occurring and Synthetic Biomaterials. *Host Response to Biomater.* 151–187 (2015). doi:10.1016/B978-0-12-800196-7.00008-6

145. Lubin, R. *et al.* Analysis of p53 antibodies in patients with various cancers define B-cell epitopes of human p53: distribution on primary structure and exposure on protein surface. *Cancer Res.* **53,** 5872–6 (1993).

146. Furthmayr, H., Beil, W. & Timpl, R. Different antigenic determinants in the polypeptide chains of human collagen. *FEBS Lett.* **12,** 341–344 (1971).

147. Benjamin, D. C. *et al. THE ANTIGENIC STRUCTURE OF PROTEINS: A REAPPRAISAL*. (1984).

148. Excellence, N. I. for H. and C. *Sorafenib for the treatment of advanced hepatocellular carcinoma*. (2010).

149. Mackay, I. R. Topics in Review: Tolerance and autoimmunity. *Western Journal of Medicine* **174,** 118 (2001).

150. Dalessandri, T. & Strid, J. Beneficial autoimmunity at body surfaces - immune surveillance and rapid type 2 immunity regulate tissue homeostasis and cancer. *Front. Immunol.* **5,** 347 (2014).

151. Pardoll, D. Does the immune system see tumors as foreign or self? *Annu. Rev. Immunol.* **21,** 807–839 (2003).

152. Dranoff, G. Cytokines in cancer pathogenesis and cancer therapy. *Nat. Rev. Cancer* **4,** 11–22 (2004).

153. Fruci, D. *et al.* Major histocompatibility complex class i and tumour immuno-evasion: how to fool T cells and natural killer cells at one time. *Curr. Oncol.* **19,** 39–41 (2012).

154. Savina, A. & Amigorena, S. Phagocytosis and antigen presentation in dendritic cells. *Immunol. Rev.* **219,** 143–56 (2007).

155. Forner, A., Reig, M. & Bruix, J. Hepatocellular carcinoma. *Lancet* **391,** 1301–1314 (2018).

156. Suppiah, A. & Greenman, J. Clinical utility of anti-p53 auto-antibody: systematic review and focus on colorectal cancer. *World J. Gastroenterol.* **19,** 4651–70 (2013).

157. Malkin, D. Li-fraumeni syndrome. *Genes Cancer* **2,** 475–84 (2011).

158. Disis, M. L., Knutson, K. L., Schiffman, K., Rinn, K. & McNeel, D. G. Pre-existent immunity to the HER-2/neu oncogenic protein in patients with HER-2/neu overexpressing breast and ovarian cancer. *Breast Cancer Res. Treat.* **62,** 245–252 (2000).

159. Brass, N. *et al.* Role of amplified genes in the production of autoantibodies. *Blood* **93,** 2158–66 (1999).

160. Sahin, U. *et al.* Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. U. S. A.* **92,**

11810–3 (1995).

161. Stockert, E. *et al.* A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. *J. Exp. Med.* **187,** 1349–54 (1998).

162. Güre, A. O. *et al.* Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 4198–203 (2000).

163. Wandall, H. H. *et al.* Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. *Cancer Res.* **70,** 1306–13 (2010).

164. Bozkaya, G. *et al.* Cooperative interaction of MUC1 with the HGF/c-Met pathway during hepatocarcinogenesis. *Mol. Cancer* **11,** 64 (2012).

165. Chen, Y.-T., Gure, A. O. & Scanlan, M. J. Serological analysis of expression cDNA libraries (SEREX): an immunoscreening technique for identifying immunogenic tumor antigens. *Methods Mol. Med.* **103,** 207–16 (2005).

166. Lodish, H. *et al.* Constructing DNA Libraries with λ Phage and Other Cloning Vectors. (2000).

167. Cottrell, T. R., Hall, J. C., Rosen, A. & Casciola-Rosen, L. Identification of novel autoantigens by a triangulation approach. *J. Immunol. Methods* **385,** 35–44 (2012).

168. Chapman, C. J. *et al.* EarlyCDT®-Lung test: improved clinical utility through additional autoantibody assays. *Tumour Biol.* **33,** 1319–26 (2012).

169. Doseeva, V., Colpitts, T., Gao, G., Woodcock, J. & Knezevic, V. Performance of a multiplexed dual analyte immunoassay for the early detection of non-small cell lung cancer. *J. Transl. Med.* **13,** 55 (2015).

170. Middleton, C. H. *et al.* Serum Autoantibody Measurement for the Detection of Hepatocellular Carcinoma. *PLoS One* **9,** e103867 (2014).

171. Chen, Y. *et al.* Autoantibodies to tumor-associated antigens combined with abnormal alpha-fetoprotein enhance immunodiagnosis of hepatocellular carcinoma. *Cancer Lett.* **289,** 32–39 (2010).

172. Dai, L. *et al.* Using immunomic approach to enhance tumor-associated autoantibody detection in diagnosis of hepatocellular carcinoma. *Clin. Immunol.* **152,** 127–139 (2014).

173. Anderson, K. S. & LaBaer, J. The Sentinel Within: Exploiting the Immune System for Cancer Biomarkers. *J. Proteome Res.* **4,** 1123–1133 (2005).

174. Wallstrom, G., Anderson, K. S. & LaBaer, J. Biomarker discovery for heterogeneous diseases. *Cancer Epidemiol. Biomarkers Prev.* **22,** 747–55 (2013).

175. Menon, U. *et al.* Risk Algorithm Using Serial Biomarker Measurements Doubles the Number of Screen-Detected Cancers Compared With a Single-Threshold Rule in the United Kingdom Collaborative Trial of Ovarian Cancer Screening. *J. Clin. Oncol.* **33,** 2062–71 (2015).

176. Ruopp, M. D., Perkins, N. J., Whitcomb, B. W. & Schisterman, E. F. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom. J.* **50,** 419–30 (2008).

177. Middleton, C. H. Recombinant tumor-associated antigen production and autoantibody detection in hepatocellular carcinoma. (University of Nottingham, 2013).

178. Larsen, J., Lund, O. & Nielsen, M. Improved method for predicting linear B-cell epitopes. *Immunome Res.* **2,** 2 (2006).

179. Macdonald, I. K. *et al.* Development and Validation of a High Throughput System for Discovery of Antigens for Autoantibody Detection. *PLoS One* **7,** e40759 (2012).

180. Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M. & Panda, A. K. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb. Cell Fact.* **14,** 41 (2015).

181. Yau, W. Y. *et al.* Autoantibody recognition of an N-terminal epitope of hnRNP L marks the risk for developing HBV-related hepatocellular carcinoma. *J. Proteomics* **94,** 346–358 (2013).

182. Li, L., Chen, S. H., Yu, C. H., Li, Y. M. & Wang, S. Q. Identification of hepatocellular-carcinoma-associated antigens and autoantibodies by serological proteome analysis combined with protein microarray. *J. Proteome Res.* **7,** 611–620 (2008).

183. Xu, Y. *et al.* Expression of Ku86 and presence of Ku86 Antibody as biomarkers of hepatitis B virus related hepatocellular carcinoma. *Dig. Dis. Sci.* **59,** 614–622 (2014).

184. Nomura, F. *et al.* Serum anti-Ku86 is a potential biomarker for early detection of hepatitis C virus-related hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* **421,** 837–843 (2012).

185. Zhang, J. Autoantibodies against glucose-regulated protein 78 as serological diagnostic biomarkers in hepatocellular carcinoma. *International Journal of Oncology* (2012). doi:10.3892/ijo.2012.1515

186. Wang, K. *et al.* Immunogenicity of Ra1A and its tissue-specific expression in hepatocellular carcinoma. *Int. J. Immunopathol. Pharmacol.* **22,** 735–43

187. Liu, H. *et al.* Screening of autoantibodies as potential biomarkers for hepatocellular carcinoma by using T7 phase display system. *Cancer Epidemiol.* **36,** 82–8 (2012).

188. Zhou, S.-F. *et al.* Identification of HCC-22-5 tumor-associated antigen and antibody response in patients. *Clin. Chim. Acta.* **366,** 274–280 (2006).

189. Looi, K. S. *et al.* Using proteomic approach to identify tumor-associated antigens as markers in hepatocellular carcinoma. *J. Proteome Res.* **7,** 4004–4012 (2008).

190. Raedle, J. *et al.* Clinical evaluation of autoantibodies to p53 protein in patients with chronic liver disease and hepatocellular carcinoma. *Eur. J. Cancer* **34,** 1198–1203 (1998).

191. Zhang, J. Y., Chan, E. K., Peng, X. X. & Tan, E. M. A novel cytoplasmic

protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma. *J. Exp. Med.* **189,** 1101–10 (1999).

192. Liu, M. *et al.* A cancer-related protein 14-3-3?? is a potential tumor-associated antigen in immunodiagnosis of hepatocellular carcinoma. *Tumor Biology* 1–10 (2014). doi:10.1007/s13277-013-1555-8

193. Liu, M. *et al.* Autoantibody response to murine double minute 2 protein in immunodiagnosis of hepatocellular carcinoma. *J. Immunol. Res.* **2014,** 906532 (2014).

194. Le Naour, F. *et al.* A distinct repertoire of autoantibodies in hepatocellular carcinoma identified by proteomic analysis. *Mol. Cell. Proteomics* **1,** 197–203 (2002).

195. Wang, K. *et al.* Identification of tumor-associated antigens by using SEREX in hepatocellular carcinoma. *Cancer Lett.* **281,** 144–150 (2009).

196. Korangy, F. *et al.* Spontaneous tumor-specific humoral and cellular immune responses to NY-ESO-1 in hepatocellular carcinoma. *Clin. Cancer Res.* **10,** 4332–4341 (2004).

197. Himoto, T. *et al.* Analyses of autoantibodies against tumor-associated antigens in patients with hepatocellular carcinoma. *Int J Oncol* **27,** 1079–1085 (2005).

198. Wang, Y. *et al.* Large scale identification of human hepatocellular carcinoma-associated antigens by autoantibodies. *J. Immunol.* **169,** 1102–1109 (2002).

199. Deininger, M. H., Meyermann, R. & Schluesener, H. J. The allograft inflammatory factor-1 family of proteins. *FEBS Lett.* **514,** 115–21 (2002).

200. Schmitz, K. J. *et al.* AKR1B10 expression is associated with less aggressive hepatocellular carcinoma: a clinicopathological study of 168 cases. *Liver Int.* **31,** 810–6 (2011).

201. Zhao, P. *et al.* Annexin II promotes invasion and migration of human hepatocellular carcinoma cells in vitro via its interaction with HAb18G/CD147. *Cancer Sci.* **101,** 387–95 (2010).

202. Fernández-Irigoyen, J. *et al.* Oxidation of specific methionine and tryptophan residues of apolipoprotein A-I in hepatocarcinogenesis. *Proteomics* **5,** 4964–72 (2005).

203. Yang, L., Si, X. & Wang, W. Overexpression of bcl-2 protects hepatoma cell line HCC-9204 from ethanol-induced apoptosis. *Chin. Med. J. (Engl).* **115,** 8–11 (2002).

204. Mima, K. *et al.* CD44s regulates the TGF-β-mediated mesenchymal phenotype and is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Res.* **72,** 3414–23 (2012).

205. Wang, Y. *et al.* JAB1 and phospho-Ser10 p27 expression profile determine human hepatocellular carcinoma prognosis. *J. Cancer Res. Clin. Oncol.* **140,** 969–78 (2014).

206. Spadaro, A. *et al. Serum chromogranin-A in hepatocellular carcinoma: diagnostic utility and limits. World journal of gastroenterology : WJG* **11,** (2005).

207. John, K., Wielgosz, S., Schulze-Osthoff, K., Bantel, H. & Hass, R. Increased plasma levels of CK-18 as potential cell death biomarker in patients with HELLP syndrome. *Cell Death Dis.* **4,** e886 (2013).

208. Yoneda, N. *et al.* Epidermal growth factor induces cytokeratin 19 expression accompanied by increased growth abilities in human hepatocellular carcinoma. *Lab. Invest.* **91,** 262–72 (2011).

209. Liu, H., Dong, J. (Huijia), Zhang, L., Robertson, K. & Liu, C. Hep Par 1 antigen (CPS1) expression in liver cancer is regulated by DNA methylation. *FASEB J* **23,** 117.12- (2009).

210. Chang, P.-C. *et al.* DDX3, a DEAD box RNA helicase, is deregulated in hepatitis virus-associated hepatocellular carcinoma and is involved in cell growth control. *Oncogene* **25,** 1991–2003 (2006).

211. Nakamura, J. *et al.* Overexpression of eukaryotic elongation factor eEF2 in gastrointestinal cancers and its involvement in G2/M progression in the cell cycle. *Int. J. Oncol.* **34,** 1181–1189 (2009).

212. Yamashita, T. *et al.* EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. *Gastroenterology* **136,** 1012–24 (2009).

213. Okuda, H., Obata, H., Nakanishi, T., Furukawa, R. & Hashimoto, E. Production of abnormal prothrombin (des-γ-carboxy prothrombin) by hepatocellular carcinoma. *J. Hepatol.* **4,** 357–363 (1987).

214. Wang, K. *et al.* Alpha-1-fucosidase as a prognostic indicator for hepatocellular carcinoma following hepatectomy: a large-scale, long-term study. *Br. J. Cancer* **110,** 1811–9 (2014).

215. Long, J. *et al.* Expression level of glutamine synthetase is increased in hepatocellular carcinoma and liver tissue with cirrhosis and chronic hepatitis B. *Hepatol. Int.* **5,** 698–706 (2011).

216. Sun, Y. *et al.* Increased Golgi protein 73 expression in hepatocellular carcinoma tissue correlates with tumor aggression but not survival. *J. Gastroenterol. Hepatol.* **26,** 1207–12 (2011).

217. Yamagamim, H. *et al.* Serum concentrations of human hepatocyte growth factor is a useful indicator for predicting the occurrence of hepatocellular carcinomas in C-viral chronic liver diseases. *Cancer* **95,** 824–34 (2002).

218. Shilo, A. *et al.* Splicing factor hnRNP A2 activates the Ras-MAPK-ERK pathway by controlling A-Raf splicing in hepatocellular carcinoma development. *RNA* **20,** 505–15 (2014).

219. Scaggiante, B. *et al.* Novel hepatocellular carcinoma molecules with prognostic and therapeutic potentials. *World J. Gastroenterol.* **20,** 1268–88 (2014).

220. Fu, Y. *et al.* Expression of HSPA2 in human hepatocellular carcinoma and its clinical significance. *Tumour Biol.* **35,** 11283–7 (2014).

221. Yang, Z. *et al.* Upregulation of heat shock proteins (HSPA12A, HSP90B1, HSPA4, HSPA5 and HSPA6) in tumour tissues is associated with poor outcomes from HBV-related early-stage hepatocellular carcinoma. *Int. J. Med. Sci.* **12,** 256–63 (2015).

222. Johnson, C. *et al.* Interleukin-6 and its receptor, key players in hepatobiliary inflammation and cancer. *Translational Gastrointestinal Cancer* **1,** 58–70 (2011).

223. Tang, K. H. *et al.* CD133(+) liver tumor-initiating cells promote tumor angiogenesis, growth, and self-renewal through neurotensin/interleukin-8/CXCL1 signaling. *Hepatology* **55,** 807–20 (2012).

224. Starmann, J. *et al.* Gene expression profiling unravels cancer-related hepatic molecular signatures in steatohepatitis but not in steatosis. *PLoS One* **7,** e46584 (2012).

225. Sun, S., Xu, M. Z., Poon, R. T., Day, P. J. & Luk, J. M. Circulating Lamin B1 (LMNB1) biomarker detects early stages of liver cancer in patients. *J. Proteome Res.* **9,** 70–8 (2010).

226. Zhang, H.-G. *et al.* Specific CD8(+)T cell responses to HLA-A2 restricted MAGE-A3 p271-279 peptide in hepatocellular carcinoma patients without vaccination. *Cancer Immunol. Immunother.* **56,** 1945–54 (2007).

227. Meng, X., Franklin, D. A., Dong, J. & Zhang, Y. MDM2-p53 pathway in hepatocellular carcinoma. *Cancer Res.* **74,** 7161–7 (2014).

228. Chen, R. *et al.* The significance of MMP-9 over MMP-2 in HCC invasiveness and recurrence of hepatocellular carcinoma after curative resection. *Ann. Surg. Oncol.* **19 Suppl 3,** S375-84 (2012).

229. Yu, L. *et al.* Mesothelin as a potential therapeutic target in human cholangiocarcinoma. *J. Cancer* **1,** 141–9 (2010).

230. Liu, X. *et al.* Nucleophosmin (NPM1/B23) interacts with activating transcription factor 5 (ATF5) protein and promotes proteasome- and caspase-dependent ATF5 degradation in hepatocellular carcinoma cells. *J. Biol. Chem.* **287,** 19599–609 (2012).

231. Xu, Y.-F. *et al.* PEBP1 downregulation is associated to poor prognosis in HCC related to hepatitis B infection. *J. Hepatol.* **53,** 872–9 (2010).

232. Yeh, Y.-T., Lee, K.-T., Tsai, C.-J., Chen, Y.-J. & Wang, S.-N. Prolactin promotes hepatocellular carcinoma through Janus kinase 2. *World J. Surg.* **36,** 1128–35 (2012).

233. Yamaguchi, M. The Regucalcin Gene is a Key in the Therapy of Hepatocellular Carcinoma. *J. Cancer Sci. Ther.*

234. Oikawa, T. *et al.* Sal-like protein 4 (SALL4), a stem cell biomarker in liver cancers. *Hepatology* **57,** 1469–83 (2013).

235. Shin, H. D. *et al.* SPP1 polymorphisms associated with HBV clearance and HCC occurrence. *Int. J. Epidemiol.* **36,** 1001–8 (2007).

236. Dillon, S. T., Bhasin, M. K., Feng, X., Koh, D. W. & Daoud, S. S. Quantitative proteomic analysis in HCV-induced HCC reveals sets of proteins with potential significance for racial disparity. *J. Transl. Med.* **11,** 239 (2013).

237. Giannelli, G., Mazzocca, A., Fransvea, E., Lahn, M. & Antonaci, S. Inhibiting TGF-β signaling in hepatocellular carcinoma. *Biochim. Biophys. Acta* **1815,** 214–23 (2011).

238. Satelli, A. & Li, S. Vimentin in cancer and its potential as a molecular target for cancer therapy. *Cell. Mol. Life Sci.* **68,** 3033–46 (2011).

239. Huang, X.-Y. *et al.* αB-crystallin complexes with 14-3-3ζ to induce epithelial-mesenchymal transition and resistance to sorafenib in hepatocellular carcinoma. *Hepatology* **57,** 2235–47 (2013).

240. Hoshino, I. *et al.* Panel of autoantibodies against multiple tumor-associated antigens for detecting gastric cancer. *Cancer Sci.* **108,** 308–315 (2017).

241. Lam, S. *et al.* EarlyCDT-Lung: an immunobiomarker test as an aid to early detection of lung cancer. *Cancer Prev. Res. (Phila).* **4,** 1126–34 (2011).

242. Chapman, C. *et al.* Autoantibodies in breast cancer: their use as an aid to early diagnosis. *Ann. Oncol.* **18,** 868–73 (2007).

243. Jung, S.-H. & Young, S. S. Power and sample size calculation for microarray studies. *J. Biopharm. Stat.* **22,** 30–42 (2012).

244. Kushner, S. R. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. *Gene* **100,** 195–199 (1991).

245. Pencina, M. J., D'Agostino, R. B., D'Agostino, R. B. & Vasan, R. S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat. Med.* **27,** 157–172 (2008).

246. Chan, S. L. *et al.* Performance of serum α-fetoprotein levels in the diagnosis of hepatocellular carcinoma in patients with a hepatic mass. *HPB* **16,** 366–372 (2014).

247. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 1760–5 (2009).

248. Yao, Y. *et al.* Potential application of non-small cell lung cancer-associated autoantibodies to early cancer diagnosis. *Biochem. Biophys. Res. Commun.* **423,** 613–619 (2012).

249. Hong, Y. & Huang, J. Autoantibodies against tumor-associated antigens for detection of hepatocellular carcinoma. *World J. Hepatol.* **7,** 1581–5 (2015).

250. Zaenker, P. & Ziman, M. R. Serologic autoantibodies as diagnostic cancer biomarkers--a review. *Cancer Epidemiol. Biomarkers Prev.* **22,** 2161–81 (2013).

251. Middleton, C. H. *et al.* Serum autoantibody measurement for the detection of hepatocellular carcinoma. *PLoS One* **9,** e103867 (2014).

252. Chan, S. L. *et al.* Performance of serum α-fetoprotein levels in the diagnosis of hepatocellular carcinoma in patients with a hepatic mass. *HPB* **16,** 366–372 (2014).

253. Ahn, D. G. *et al.* Feasibility of α-fetoprotein as a diagnostic tool for hepatocellular carcinoma in Korea. *Korean J. Intern. Med.* **31,** 46–53 (2016).

254. Sanai, F. M. *et al.* Assessment of Alpha-Fetoprotein in the Diagnosis of Hepatocellular Carcinoma in Middle Eastern Patients. *Dig. Dis. Sci.* **55,** 3568–3575 (2010).

255. Jang, E. S. *et al.* Diagnostic Performance of Alpha-Fetoprotein, Protein Induced by Vitamin K Absence, Osteopontin, Dickkopf-1 and Its Combinations for Hepatocellular Carcinoma. *PLoS One* **11,** e0151069 (2016).

256. Abbasi, A., Bhutto, A. R., Butt, N. & Munir, S. M. Corelation of serum alpha fetoprotein and tumor size in hepatocellular carcinoma. *J. Pak. Med. Assoc.* **62,** 33–6 (2012).

257. Reibnegger, G. & Schrabmair, W. Optimum binary cut-off threshold of a diagnostic test: comparison of different methods using Monte Carlo technique. *BMC Med. Inform. Decis. Mak.* **14,** 99 (2014).

258. Draminski, M. *et al.* Monte Carlo feature selection for supervised classification. *Bioinformatics* **24,** 110–117 (2008).

259. Wallstrom, G., Anderson, K. S. & LaBaer, J. Biomarker discovery for heterogeneous diseases. *Cancer Epidemiol. Biomarkers Prev.* **22,** 747–55 (2013).

260. Chen, J. *et al.* Hepatic resection after transarterial chemoembolization increases overall survival in large/multifocal hepatocellular carcinoma: a retrospective cohort study. *Oncotarget* **8,** 408–417 (2017).

261. Tao, X., Wang, N. & Qin, W. Gut Microbiota and Hepatocellular Carcinoma. *Gastrointest. tumors* **2,** 33–40 (2015).

262. Heimbach, J. K. *et al.* AASLD Guidelines for the Treatment of Hepatocellular Carcinoma Guiding Principles and Objectives GUIDING PRINCIPLES. **67,** (2018).

263. Yoshida, H. *et al.* Bacterial infection in cirrhosis, with and without hepatocellular carcinoma. *Am. J. Gastroenterol.* **88,** 2067–71 (1993).

264. Papp, M. *et al.* Presence of anti-microbial antibodies in liver cirrhosis--a tell-tale sign of compromised immunity? *PLoS One* **5,** e12957 (2010).

265. Harper, S. & Speicher, D. W. Purification of proteins fused to glutathione S-transferase. *Methods Mol. Biol.* **681,** 259–80 (2011).

266. Tebeje, B. M., Harvie, M., You, H., Loukas, A. & McManus, D. P. Schistosomiasis vaccines: where do we stand? *Parasit. Vectors* **9,** 528 (2016).

267. Chou, R. *et al.* Imaging Techniques for the Diagnosis of Hepatocellular Carcinoma. *Ann. Intern. Med.* **162,** 697 (2015).

268. Zhou, J. *et al.* Guidelines for Diagnosis and Treatment of Primary Liver Cancer in China (2017 Edition). *Liver Cancer* 1–26 (2018). doi:10.1159/000488035

269. Chen, J.-G. *et al. Screening for liver cancer: results of a randomised controlled trial in Qidong, China*. *204 Journal of Medical Screening* **10,** (2003).

270. Singal, A. G. *et al.* Effectiveness of Hepatocellular Carcinoma

Surveillance in Patients with Cirrhosis. *Cancer Epidemiol. Biomarkers Prev.* **21,** 793–799 (2012).

271. Chou, R. *et al.* Imaging Techniques for the Diagnosis of Hepatocellular Carcinoma. *Ann. Intern. Med.* **162,** 697 (2015).

272. Šimundić, A.-M. Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC* **19,** 203–11 (2009).

273. Ma, X. *et al.* Risk prediction models for hepatocellular carcinoma in different populations. *Chin. J. Cancer Res.* **28,** 150–60 (2016).

274. Papatheodoridis, G. *et al.* PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. *J. Hepatol.* **64,** 800–806 (2016).

275. Lou, J., Zhang, L., Lv, S., Zhang, C. & Jiang, S. Biomarkers for Hepatocellular Carcinoma. *Biomark. Cancer* **9,** 1–9 (2017).

276. Cohen, J. D. *et al.* Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 10202–10207 (2017).

277. Khan, J., Lieberman, J. A. & Lockwood, C. M. Variability in, variability out: best practice recommendations to standardize pre-analytical variables in the detection of circulating and tissue microRNAs. *Clin. Chem. Lab. Med.* **55,** 608–621 (2017).

278. Sanchez-Carbayo, M. Antibody arrays: technical considerations and clinical applications in cancer. *Clin. Chem.* **52,** 1651–9 (2006).

279. Lim, C. T. & Zhang, Y. Bead-based microfluidic immunoassays: The next generation. *Biosens. Bioelectron.* **22,** 1197–1204 (2007).

280. Jia, Z. *et al.* Detection of Lung Cancer: Concomitant Volatile Organic Compounds and Metabolomic Profiling of Six Cancer Cell Lines of Different Histological Origins. *ACS omega* **3,** 5131–5140 (2018).

281. Guo, W., Tan, H. Y., Wang, N., Wang, X. & Feng, Y. Deciphering hepatocellular carcinoma through metabolomics: from biomarker discovery to therapy evaluation. *Cancer Manag. Res.* **10,** 715–734 (2018).

282. Luo, P. *et al.* A Large-scale, multicenter serum metabolite biomarker identification study for the early detection of hepatocellular carcinoma. *Hepatology* **67,** 662–675 (2018).

283. Pillai, R. *et al.* Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed, paraffin-embedded specimens. *J. Mol. Diagn.* **13,** 48–56 (2011).

284. Yang, N. *et al.* The characteristics of ctDNA reveal the high complexity in matching the corresponding tumor tissues. *BMC Cancer* **18,** 319 (2018).

Chapter 11. Appendices

11.1 Medias

Media/Solution	Manufacturer	Description/Final concentration	Unit quantity per 1L
LB Agar	Fisher BPE1425	LB Agar (Miller)	40g
LB Broth	Fisher BPE1426	LB Broth (Miller)	25g
Turbo Broth Overnight expression medium (TBOnex-P)	Novagen 71491-5	Overnight Express Instant TB Media	60g
	Fisher 158920025	1% Glycerol	10ml
	Sigma S0751	NaH2PO4	2.76g
	Sigma S7907	Na2HPO4	10.93g
2xYT	Fisher 1279-7099	Tryptone	16g
	Fisher BPE1422	Yeast Extract	10g
	Fisher S/3160/63	NaCl	5g

11.2 Purification buffers

Buffer	Manufacturer	Description/Final concentration	Unit quantity per 1L
IB lysis buffer	Fisher BPE178	6M GuHCl	573.18g
	From 1M stock	50mM Tris-HCl pH8	50ml
	Sigma 56749	20% Imidazole	1.36g
	Fisher S/3160/63	0.5M NaCl	29.22g
	Fisher 158920010	10% Glycerol	100ml
	Promega H5142	1% Triton	10ml
	From 0.5M stock	1mM TCEP	2ml
Soluble lysis buffer	Merck 71456	BugBuster Master Mix	990ml
	Merck 539137	Protease Inhibitor Cocktail Set V	10ml
Soluble binding buffer	Sigma 56749	20mM Imidazole	1.36g
	Fisher 158920010	10% Glycerol	100ml
	From 1M stock	50mM Tris solution pH 8	50ml
	Fisher S/3160/63	0.5M NaCl	29.22g
Urea wash buffer	Fisher U/0500/53	8M Urea	480.48g
	From 1M stock	50mM Tris-HCl pH8	50ml
	Fisher 158920010	10% Glycerol	100ml
	Fisher S/3160/63	0.5M NaCl	29.22g
	Sigma 56749	20mM Imidazole	1.36g
Guanidine wash buffer	Sigma G3272	6M GuHCl	573.18g
	From 1M stock	50mM Tris-HCl pH8	50ml

	Fisher S/3160/63	0.5M NaCl	29.22g
	Sigma 56749	20mM Imidazole	1.36g
	Fisher 158920010	10% Glycerol	100ml
	From 0.5M stock	1mM TCEP	2ml
Urea refolding buffer	Fisher U/0500/53	1M Urea	60.06g
	From 1M Stock	50mM Tris-HCl pH8	50ml
	Fisher S/3160/63	0.2M NaCl	11.69g
	Fisher 120000250	2mM Glutathione Reduced	614mg
	Fisher 320225000	0.02mM Glutathione Oxidized	12.2mg
	Sigma 56749	20mM Imidazole	1.36g
	Fisher 158920010	10% Glycerol	100ml
	Fisher BPE337	0.005% Tween	0.05ml

Conditioning buffer	Fisher U/0500/53	1M Urea	60.06g
	From 1M stock	50mM Tris-HCl pH8	50ml
	Fisher S/3160/63	0.2M NaCl	11.69g
	Fisher 158920010	10% Glycerol	100ml
	Sigma 56749	20mM Imidazole	1.36g
Elution buffer	Fisher U/0500/53	1M Urea	60.06g
	From 1M stock	50mM Tris-HCl pH8	50ml
	Fisher S/3160/63	0.2M NaCl	11.69g
	Fisher 158920010	10% Glycerol	100ml
	Sigma 56749	250mM Imidazole	17.02g
Dialysis buffer 2	Fisher U/0500/53	1M Urea	60.06g
	From 1M stock	50mM Tris solution pH8	50ml
	Fisher S/3160/63	0.2M NaCl	11.69g
	Fisher 158920010	10% Glycerol	100ml
Urea binding buffer	buffer composition is propriety information controlled by Oncimmune Ltd™		
IB solubilisation buffer			
Guanidine binding buffer			

Urea elution buffer	
Urea column scrubbing buffer	
4M dialysis buffer	
2M dialysis buffer	
1M dialysis buffer	
Urea dialysis buffer 1	
Urea dialysis buffer 2	

11.3 Stock solutions

Buffer	Manufacturer	Description/Final concentration	Unit quantity per 1L
1M Stock Tris-HCl	Fisher 228030010	Tris HCl	88.8g
	Sigma T6066	Trizma Base	53g
0.5M Tris(2-carboxyethyl)phosphine HCl (TCEP)	Sigma 646547	0.5M TCEP	143.3g
TBST	From 10 x Stock	1 x TBS	100ml
	Fisher BPE337	0.1% Tween 20	1ml
10xTBS	Fisher S/3160/63	1.5M NaCl	87.6g
	Fisher 228030010	Tris HCl (M.W. 157.60)	60.35g
	Sigma T6066M	Trizma Base (M.W. 121.14)	14.20g

11.4 Assay buffers

Buffer	Manufacturer	Description/Final concentration	Unit quantity per 1L
Assay coating buffer	buffer composition is propriety information controlled by Oncimmune Ltd™		
Assay blocking buffer			
Assay wash buffer			

11.5 Copyright permissions

Figure /Table	License number	License date	License provider	Authors	Publication
Figure 3	4434130228918	Sep 22, 2018	Karger Publishers	Ramakrishna Gayatri, Rastogi Archana, Trehanpati Nirupama, et al	From Cirrhosis to Hepatocellular Carcinoma: New Molecular Insights on Inflammation and Cellular Senescence. Liver Cancer 2013;2:367-383
Figure 6	4434230721632	Sep 22, 2018	Elsevier	Forner, Reig & Bruix	Hepatocellular carcinoma, The Lancet. 2018;391:1301-1314