MOLECULAR ANALYSIS OF CIRCULATING TUMOUR CELLS IN PANCREATIC CANCER

Thomas A R Hanna

Department of Molecular and Clinical Cancer Medicine

Institute of Translational Medicine

University of Liverpool

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

May 2019

DECLARATION

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration, except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university for any degree or other qualification.

ACKNOWLEDGEMENTS

Firstly, I would like to give special appreciation to my supervisors Professor William Greenhalf, Professor Paula Ghaneh and Professor John Neoptolemos who have generously given their expertise, mentorship and friendship throughout.

I would like to thank my colleagues at the Pancreas Biomedical Research Unit (PBRU) for providing support, assistance and companionship. Particularly those working closely with me on EUROPAC including Miss Sara Harrison, Dr. Li Yan, Mr. James Nicholson, Mrs Andrea Sheel, Mrs Marianne Johnson, in the laboratory Miss Hollie Pfual, and Miss Katie Bullock and to Dr. Anthony Evans for assistance with formatting the thesis.

I am grateful to Dr Nick Bryan and Professor John Hunt from the Department of Clinical Engineering for their creativity and hard work and to Mr Damian Bond and Dr Chris Stanley at Prokyma for their collaboration and support.

I am extremely grateful for the funding received from the National Institute for Health Research both through the invention for innovation stream and via PBRU.

I must give thanks to all the patients who participated so enthusiastically in the study whilst going through such challenging personal circumstances themselves and to the chemotherapy nurses who allowed me to work around them whilst being so busy.

Finally, I would like to thank my wife Candice, our two boys Oliver and Otis, and the rest of my friends and family for their unconditional support whilst studying.

LIST OF ABBREVIATIONS

ARMS	Amplification Refractory Mutation System
BSA	Bovine Serum Albumin
CBD	Common Bile Duct
CK	Cytokeratin
CP	Chronic Pancreatitis
CS	Cellsieve™
CT	Computed Tomography
CTC	Circulating Tumour Cells
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial Mesenchymal Transition
ERCP	Endoscopic Retrograde Cholangiopancreatography
EUROPAC	European Registry Of Hereditary Pancreatitis And Familial Pancreatic Cancer
EUS	Endoscopic Ultrasound
FAMMM	Familial Multiple Mole Melanoma
FBS	Foetal Bovine Serum
FDA	Food And Drug Administration
FFPE	Formalin Fixed Paraffin Embedded
FNA	Fine-Needle Aspiration
FOLFIRI	Irinotecan And Fluorouracil
FOLFIRINOX	Oxaliplatin, Irinotecan And Fluorouracil
FPC	Familial Pancreatic Cancer
GIST	Gastro-Intestinal Stromal Tumour
HP	Hereditary Pancreatitis
IARC	International Agency For Research On Cancer
ICGC	International Cancer Genome Consortium
IGV	Intergrative Genomic Viewer
INDEL	Insertion Deletion

IPMN	Intraductal Papillary Mucinous Neoplasms
IQR	Interquartile Range
ISP	Ion Sphere™ Particles
LREC	Liverpool Research Ethics Committee
MACS	Magnetic Activated Cell Sorting System
MAF	Minor Allele Frequency
MCN	Mucinous Cystic Neoplasm
MREC	Medical Research Ethics Committee
MRI	Magnetic Resonance Imaging
NCI	National Cancer Institute
NGS	Next Generation Sequencing
NHS	National Health Service
OC	OncoQuik®
OS	Overall Survival
PBMC	Peripheral Blood Mononuclear Cell
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDAC	Pancreatic Ductal Adenocarcinoma
PEP	Post-Ercp Acute Pancreatitis
PGM	Personal Genome Machine
PJS	Peutz-Jeghers Syndrome
PNET	Pancreatic Neuroendocrine Tumour
RBC	Red Blood Cells
RLUH	Royal Liverpool University Hospital
ROC	Receiver Operating Characteristic
RS	Rosettesep™
SD	Standard Deviation
SIFT	Sorting Intolerant From Tolerant
SNP	Single Nucleotide Polymorphism
SOP	Standard Operating Procedure
SSECRETIN	Secretin Stimulated Endoscopic Collection Of Duodenal Aspirate For Analysis Of Molecular Markers For The Early Detection Of Pancreatic Cancer Study
TCB	True Cut Biopsy

VEGF	Vascular Endothelial Growth Factor
WBC	White Blood Cells
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Pancreatic Ductal Adenocarcinoma (PDAC) is becoming a global health crisis. In the Western countries death rates are stable or decreasing for most forms of cancer, largely due to advances in treatment. However, for PDAC increasing incidence more than compensates for any tiny improvement in survival and so despite accounting for less than 5% of all newly diagnosed cancers, PDAC has grown to become the 4th largest cancer killer(Siegel, Miller, & Jemal, 2015) and by 2030 is projected to become the 2nd(Rahib et al., 2014). The worldwide incidence of pancreatic cancer is projected to be around 420,000 by 2020 with an associated mortality of 410,000(Ferlay et al., 2015).

The most favourable prognosis is associated with early PDAC where chemotherapy with gemcitabine plus capecitabine following surgical resection of the primary tumour giving a 5-year survival of 28%(Neoptolemos et al., 2018). FOLFIRONOX (oxaliplatin, irinotecan, leucovorin, and 5FU) has also shown to be superior to gemcitabine alone in the adjuvant setting though 5-year survival data is not yet available(Conroy et al., 2018). However, either because early PDAC is largely asymptomatic or because PDAC metastasises rapidly less than 20% of patients are suitable for surgical resection of their tumour at the time of diagnosis(Hackert, Schneider, & Büchler, 2015). For advanced disease gemcitabine has been shown to be slightly superiority to 5FU (Burris et al., 1997). For patients with a high performance indicator FOLFIRONOX (oxaliplatin, irinotecan, leucovorin, and 5FU) has been shown to be superior to gemcitabine in patients with advanced disease (Conroy et al., 2011). Targeting cell division using albumin-bound paclitaxel (nab-paclitaxel) combined with gemcitabine gives modest improvement in overall survival (OS) (Von Hoff et al., 2013). Presentation at an advanced stage, and treatment resistance are two major factors contributing to the poor 5-year survival of PDAC. Screening individuals at high risk of PDAC addresses the former, and targeted therapy the latter.

1.1 Screening in PDAC

The only chance of a cure is with surgical resection. Resectability criteria are largely determined on the basis of vascular involvement and outlined in the European Society of Medical Oncology (ESMO) guidelines(Ducreux et al., 2015). Identifying PDAC earlier in the natural history is therefore an essential strategy for improving outcomes. In this context screening is a process of selecting asymptomatic individuals, who are at

increased risk of PDAC for further tests with the aim of improving outcome for the individual(Wald, 2008). Risk factors for PDAC include smoking, diabetes mellitus, obesity and chronic pancreatitis (Coughlin, Calle, Patel, & Thun, 2000; Larsson et al., 2005; Malka, 2002; Michaud, 2004; Muscat, Steilman, & Wynder, n.d.). Up to 10% of PDAC have a hereditary component including ~~in~~ Familial Pancreatic Cancer (FPC), Hereditary Pancreatitis (HP) and other cancer syndromes such as Peutz-Jeghers (PJS), familial multiple mole melanoma (FAMMM) and breast ovarian cancer syndrome(Canto et al., 2013; Greenhalf et al., 2009; Hahn et al., 2003; Lj, Greenhalf et al., 2005; Miln et al., 2009; Vitoneet al., 2006). International consensus recommends screening only for specific groups with a genetic predisposition to PDAC calculated to have a >5% lifetime risk(Canto et al., 2013).

PDAC arises through well-defined precursor lesions which offer an opportunity for earlier diagnosis. Intraductal papillary mucinous neoplasms (IPMN) and mucinous cystic neoplasm (MCN) are both macroscopic cystic lesions >0.5cm visible by standard imaging: computed tomography (CT), magnetic resonance imaging (MRI) and endoscopic ultrasound (EUS) and can therefore be identified in asymptomatic individuals. IPMN lesions are classified into three morphological types main duct (MD), branch duct (BD) and mixed(Tanaka et al., 2012). There is wide reported variation in the risk of malignant transformation of the three types but MD-IPMN represents the largest risk; surgical resection is recommended, if both clinically appropriate and high risk stigma of malignancy are present(Jang et al., 2017). Conversely, the annual rate of progression to high grade dysplasia or invasive cancer from BD-IPMN lesions is much lower between estimated between 1.4-6.9%(Jang et al., 2017). In addition, the risk of malignancy associated with BD-IPMN has recently been shown by the EUROPAC group to be independent of the genetic risk, i.e. individuals with a higher inherited risk of PDAC are not at higher risk of developing BD-IPMN. As such the EUROPAC study does not support the inclusion of non-malignant pancreatic cystic lesions including BD-IPMN as positive findings on screening individuals from FPC families(Sheel et al., 2018).

MCN lesions are defined by the presence of ovarian stroma and have a relatively low, but still significant incidence of invasive carcinoma (<15%). Given the relatively young age of most of the patients surgical resection is generally recommended in surgically fit patients(Tanaka et al., 2012)

Pancreatic intraepithelial neoplasia (PanIN), the most common and well characterised precursor lesion is classified into four groups microscopicaly: PanIN-1A, PanIN-1B, PanIN-2, and PanIN-3 according to the degree of atypia (R. Hruban et al., 2001). Evidence of a PanIN to PDAC progression includes increasing prevalence of PanIN with age(de Wilde, Hruban, Maitra, & Offerhaus, 2011), cancer incidence(Cubilla & Fitzgerald, 1976), physical proximity of PanIN to PDAC in resected pancreata(Andea, Sarkar, & Adsay, n.d.; Furukawa et al., 1994) and the presence of all PanIN stages prior to tumour formation in genetically engineered mouse models of PDAC(Hingorani et al., 2003; Perez-Mancera, Guerra, Barbacid, & Tuveson, 2012). In addition, the four mutation 'mountains' of PDAC: the near ubiquitous point mutation in the *KRAS* oncogene and inactivating mutations of three tumour suppressor genes *TP53*, *SMAD4*, *and CDKN2A* are all found in PanIN lesions Figure 1-1. *KRAS* is an early event found in PanIN1 lesions(Lüttges et al., 1999) followed by *CDKN2A*(Mosaluk, Hruban, & Kern, 1997; Wilentz et al., 1998). SMAD4 and TP53 are late events occurring in PanIN-3 lesions with high risk of progression to PDAC(Wilentz et al., 2006) Figure 1-1

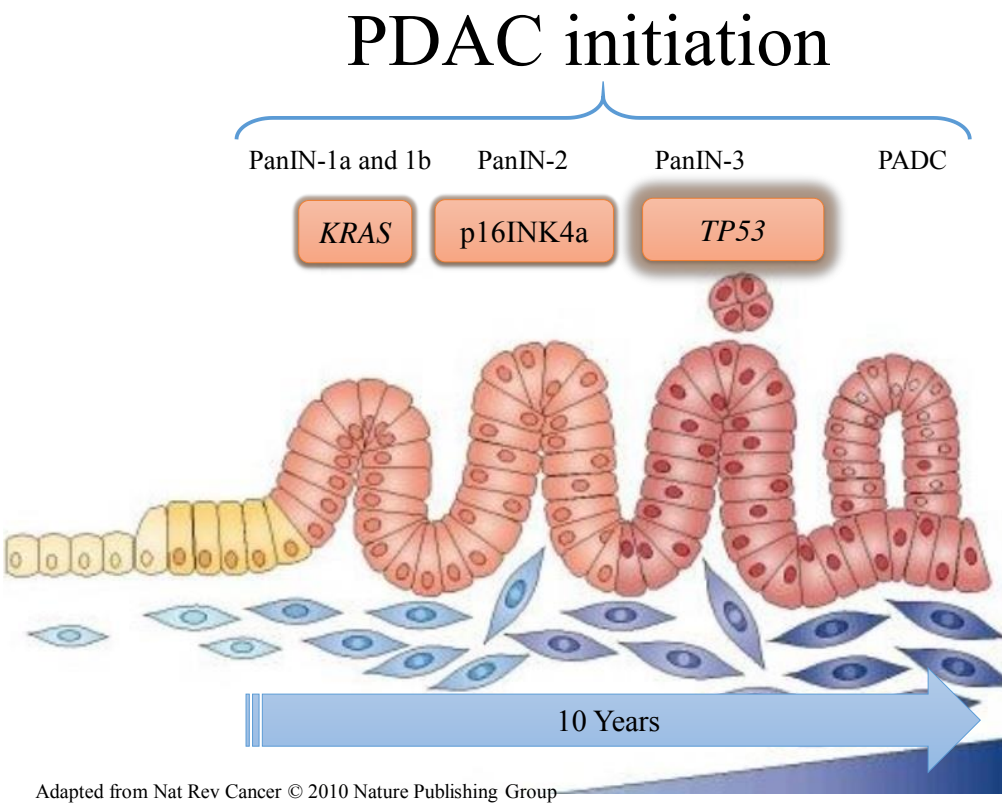


Figure 1-1 PanIN to PDAC progression model

Computational modelling based on the evolution of mutations detected by whole exome sequencing of PDAC estimates the time from the initial founder mutation (corresponding to PanIN-1) to PDAC development is 12 years, with an additional 7 years before the development of metastasis(Campbell et al., 2010; S Yachida & Iacobuzio-Donahue, 2013; Shinichi Yachida et al., 2011) potentially offering a 19-year window to achieve curative resection if the genetic markers can be identified; although this may be a considerable overestimate, it does support the view that there is a considerable window of opportunity to detect PDAC in asymptomatic individuals. PDAC therefore meets several of the screening principles (Wilson & Jungner, 1968); it is an important health problem, has an adequately understood progression from latent phase to disease, with an accepted treatment more effective earlier in the natural history. There also exists an agreed policy on who to screen. However, the requirement for a suitable and acceptable test has arguably not yet been achieved and the further work, described in chapter one is sorely needed.

The European Registry of Hereditary Pancreatitis and Familial Pancreatic Cancer (EUROPAC) is an on-going research study exploring the potential of various screening modalities including the molecular analysis of pancreatic juice in patients with FPC, HP and other cancer syndromes since 1999 (Grocock et al., 2007; Lj et al., 2005)(Yan et al., 2005). The screening protocol includes serum tests for Ca19.9, and regular imaging of the pancreas with CT, MRI and EUS to identify macroscopic precursor lesions: IPMN and MCNs. In addition, molecular analysis of pancreatic juice to detect the genetic changes in the latent period before PDAC develops is also performed by EUROPAC and other groups. The simplest method of obtaining pure pancreatic juice is by endoscopic retrograde cholangiopancreatography (ERCP) and direct cannulation of the main pancreatic duct but this causes an unacceptably high rate of post-ERCP acute pancreatitis (PEP)(Freeman ML, DiSario JA, Nelson DB, 2001). The use of prophylaxis to reduce the risk of PEP and improve the clinical utility of pancreatic juice as a biospecimen has been examined by the EUROPAC group and found to be reasonably effective in doing so(Nicholson et al., 2015). In this paper there were 13 (16.3%) cases of post-ERCP acute pancreatitis (PEP) all of which occurred in the FPC cohort (23.2%). On 21st November 2008 the use of prophylaxis (self-expelling 5F plastic pancreatic stents and rectal diclofenac) was introduced in an attempt to reduce the rate of PEP. Prior to 21st November 2008 the incidence of PEP in the FPC group was 7/34 (21%) which subsequently fell to 6/40 (15%) following the introduction of prophylaxis

p=0.0347. Overall analysis of the 56 ERCP procedures performed in the FPC group showed that the omission of prophylaxis was the only significant association with PEP p=0.0347. Prophylaxis measures were not used in patients in the HP cohort as the risk of PEP was low (no events) in this study. However, the risk of PEP even with prophylaxis was deemed still too high to justify routine screening in clinical practice. The risk of PEP may be avoided altogether if the pancreatic juice sample is collected from the duodenum and cannulation of the main pancreatic duct is avoided. This method employs the use of a synthetic gut hormone Secretin, which is administered intravenously and stimulates the production and secretion of pancreatic juice. The stimulated juice is then collected within the duodenum, avoiding pancreatic duct cannulation and hence the associated risk of pancreatitis. However, sampling juice from the duodenum introduces variables such as contamination with gut flora, mixing with duodenal derived deoxyribonucleic acid (DNA) (and dilution with alimentary tract fluid, which may interfere with the downstream analysis and reduce the sensitivity and specificity for PDAC). The Secretin Stimulated Endoscopic Collection of Duodenal Aspirate for Analysis of Molecular Markers for the Early Detection of Pancreatic Cancer study (SSECRETIN) investigated the diagnostic value of pancreatic juice collected in this manner. The study initially evaluated the supernatant and pellet of pancreatic juice. During the course of the study, the collection of whole pancreatic juice for this purpose was reported (M. Kanda et al., 2012; Mitsuro Kanda, Sadakari, et al., 2013). A minor amendment to the SSECRETIN study protocol was made permitting collection of whole pancreatic juice in addition to the pellet and supernatant in order to make a direct comparison and maximize the potential of this biospecimen. The study recruited patients with a clinical and/or radiological diagnosis of pancreatic/biliary cancer or benign diseas who are undergoing ERCP for a clinical indication, such as histological sampling, stent insertion or bile duct stone removal.

1.2 Targeted therapy

In addition to improved screening and earlier diagnosis, improving response to chemotherapy and overcoming resistance is vital. The targeted therapy revolution, in which newly discovered signalling networks found to be altered in cancer cells are specifically targeted by small molecules (Hanahan & Weinberg, 2011)(Chabner & Roberts, 2005) has improved survival in many cancer types, but not PDAC. For

example, bevacizumab (a recombinant humanized anti- Vascular Endothelial Growth Factor (VEGF) monoclonal antibody) gave improved overall survival (OS) in phase III trials in advanced colorectal(Hurwitz et al., 2004) non-small cell lung(Sandler A, Gray R, Perry MC, 2006), renal cell(Escudier et al., 2007), and breast cancer(Miller et al., 2007). Despite strong pancreas specific pre-clinical evidence(Bockhorn et al., 2003) and an encouraging phase II study(Hedy L Kindler et al., 2005) no improvement in OS was seen in a phase III study in advanced PDAC(H. L. Kindler et al., 2010). Unfortunately bevacizumab is just one example of many targeted therapies which follow this pattern(Bramhall et al., 2002; Gonçalves et al., 2012; Hedy L Kindler et al., 2011; Philip et al., 2010; Rougier et al., 2013; Van Cutsem, 2004). The problem is that all the targeted agents described above might be effective against PDAC in some patients, none of them is effective for the majority of patients and this means it is difficult to establish efficacy in a clinical trial. The genetic heterogeneity of PDAC responsible for these failures are discussed later.

Current approaches to capitalise on advances in genomic medicine to improve outcomes in PDAC include Precision-Panc; a collection of phase two studies under the umbrella of Pancreatic cancer Individualised Multi-arm Umbrella Studies (PRIMUS)(UK, 2019). PRIMUS 001 is a phase two study comparing two chemotherapy regimens FOLFOX and nab-paclitaxel versus gemcitabine and nab-paclitaxel in patients with advanced PDAC focusing on in-depth molecular profiling and biomarker discovery. PRIMUS-2 examines two neo-adjuvant regimes FOLFOX with nab-paclitaxel and gemcitabine with nab-paclitxel in resectable and borderline resectable PDAC focusing on biomarker and liquid biopsy development and opened at it's first site in Glasgow, UK in March 2019. The end date of recruitment is expected to be September 2022.

1.3 Genetic heterogeneity in PDAC

The International Cancer Genome Consortium (ICGC) set up large-scale cancer genome studies to generate a comprehensive catalogue of somatic mutations from a variety of cancer types including PDAC(Hudson et al., 2010). The studies re-confirmed the importance of the mountain mutations in PDAC; *TP53, KRAS, CDKN2A and SMAD4* (Biankin et al., 2012; Nones et al., 2014; Waddell et al., 2015) but also promisingly in 30% of patients many actionable mutations were discovered including: *ERBB2* (*HER2*)

targeted by the monoclonal antibody trastuzumab(Slamon et al., 1987) focal amplification involving CDK4/6 targeted by palbociclib(Franco., et al 2014), *PIK3CA* mutations modulated by inhibitors of the mTOR pathway(Asano et al., 2005; Bondar et al., 2002; Bruns et al., 2004; Ito et al., 2006) *MET* mutations targeted by tivantinib(Perez-Ramirez et al., 2015) and finally mutations in the *FGFR1* gene targeted with ponatinib(Gozgit et al., 2012). Despite this encouraging discovery, all these actionable mutations were identified at low individual prevalence (1-2% of patients)(Waddell et al., 2015). Such low frequency actionable mutations have significant implications for biospecimen interrogation in terms of sampling error. In addition, the marked heterogeneity has another consequence which promotes resistance to chemotherapy through clonal evolution.

1.4 Clonal Evolution

The traditional view of cancer has been that tumorigenesis is caused by the successive accumulation of a handful of mutations in oncogenes and tumour suppressor genes(Hanahan & Weinberg, 2000; Vogelstein & Kinzler, 2004). By re-considering cancer as inherently an evolutionary process, the development of chemotherapy resistance can be better understood(Greaves & Maley, 2012a)(Hanahan & Weinberg, 2011). In the evolution model the accumulation of a few mutations in key genes results in genomic instability which contribute to the formation of multiple clonal subpopulations with distinct molecular profiles and therefore distinct sensitivities to chemotherapies. These have been confirmed by sequencing from multiple topographical sites within the same primary(Gerlinger et al., 2012; Martinez et al., 2013; Sottoriva et al., 2013). Linear evolution maintains that sub-clones in the primary are formed before cancer dissemination and simply enriched in metastasis based on selective pressure(Shinichi Yachida et al., 2011). In this scenario all mutations in the metastasis would be present in the primary tumour, so called ‘trunk’ mutations. During chemotherapy massive tumour cell death of sensitive clones occurs, but resistant sub-clones of cancer cells remain: a clonal sweep(Beerenwinkel et al., 2007). These subclones no longer need to compete for the resources of the local microenvironment and rapidly multiply emerging as tumour recurrence, with a new, distinct, molecular profile. Support for the evolution model is seen by observing the degree of initial primary genetic heterogeneity is an essential determinant of progression to cancer in

pre-malignant cells (Maley et al., 2006), to the development of metastasis(Fidler, 2003; Greaves & Maley, 2012a; Gupta & Massagué, 2006) and the acquisition of treatment resistance(Korolev, Xavier, & Gore, 2014).

Branched evolution implies that additional driver mutations (branch mutations) develop after the cancer dissemination; with significant implications for personalised therapy based on primary tumour analysis. This is supported by sequencing studies of multiple metastatic sites in PDAC demonstrating in 7/10 patients branch mutations were found exclusively in the metastasis(Campbell et al., 2010). The discrepancy between primary and metastasis was considerable with one patient having 8 genetic rearrangements in all metastatic sites not present in the primary inferring one of these at least is a driver mutation(Campbell et al., 2010).

Personalised therapy based on any actionable mutations found in the primary will be futile in this scenario, where the driver of tumorigenesis has long been replaced. Theoretically, the latest dominant sub-clone harbouring the new branch mutation which is now driving tumourigenesis can be re-characterised and targeted with a different agent to which it is sensitive. However, there is currently no way of tracking the disease and its molecular profile in this way after successive clonal sweeps. This work focuses on developing such a method.

1.5 Biospecimens in PDAC

The marked genetic heterogeneity and propensity for branched clonal evolution alone place huge demands on a biospecimen for PDAC if personalised therapy is to be realised. It must accurately reflect the contemporaneous molecular composition of the tumour biology; be adequate for analysis (e.g. offers enough DNA of good enough quality for sequencing); and allow minimally invasive serial sampling to track branched clonal evolution Figure 1-2

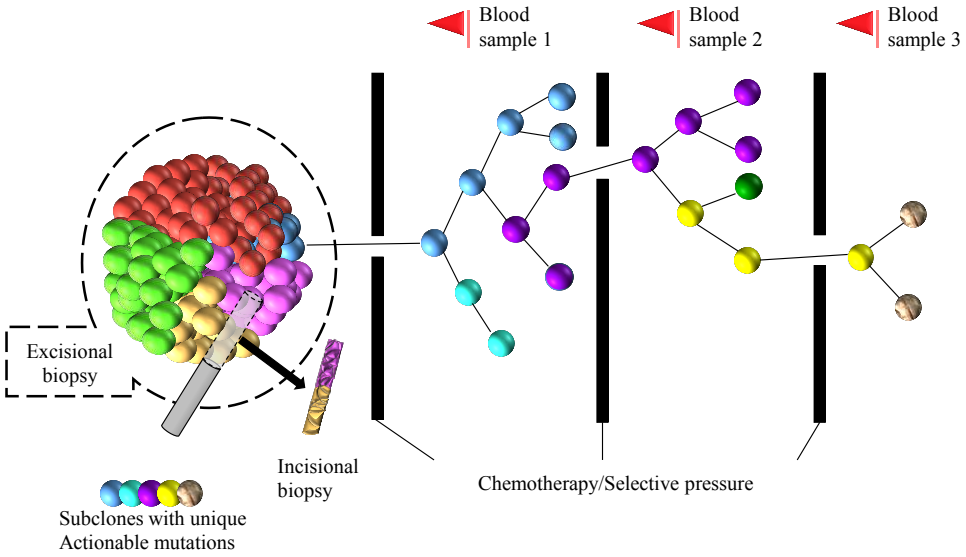


Figure 1-2 Branched evolution in PDAC and various biospecimens. Incisional biopsy samples have only a fraction of primary tumour heterogeneity. Low frequency subclones (blue) enriched in metastasis, such that excisional biopsy is unrepresentative.

1.5.1. Incisional biopsy

In breast, colorectal and ovarian cancer the majority of patients undergo surgical resection, but in PDAC, patients undergoing resection are in the minority at 17%(Speelman, Gestel, Rutten, & Hingh, 2015). As a result, in the majority of cases molecular profiling of primary PDAC is reliant on incisional rather than excisional biopsies. The pancreas however, occupies a retroperitoneal position, in close proximity to major vascular structures such that endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA) is usually the sole method of obtaining a tissue diagnosis. The aspirate is often of limited or no cellularity and inadequate for diagnosis let alone Next Generation Sequencing (NGS)(Yadav, Li, Lavery, Yadav, & Tewari, 2015). Whilst successful NGS on pancreas FNA has been reported and in fact shown good concordance with paired formalin fixed paraffin embedded (FFPE) samples from the primary tumour, this approach has not been widely replicated(Young et al., 2013). However, this problem has been somewhat mitigated by modifications to sampling method. Fine needle biopsy (FNB) involves multiple non-parallel passes through the tumour using a fork shaped needle. Recent studies have shown that EUS-FNB can now obtain sufficient tissue for targeted NGS in pancreatic adenocarcinoma in over 70% of

cases(Elhanafi et al., 2018). Indeed currently underway are large scale whole genome sequencing projects which using this methodology.

Core tissue biopsy is the gold standard incisional biopsy and often is a minimum requirement for clinical trials enrolment. EUS-guided Tru-Cut biopsies (EUS-TCB) of the pancreas was first reported in 2002(Wiersema et al., 2002), however, this is a technically difficult procedure and as a consequence improvement in diagnostic accuracy (over FNA) proved marginal in early studies(Shah et al., 2008) and EUS-TCB has not yet been adopted into routine clinical practice(Fuccio & Larghi, 2014). Regardless of feasibility, incisional biopsies are inherently limited in heterogeneous cancers only sampling a fraction of the most dominant sub-clones within the primary(Gerlinger et al., 2012), which in any case become irrelevant once branched evolution begins Figure 1-2.

1.5.2. Excisional biopsy

Excisional biopsies (resection) of the primary tumour fare little better not least because they are limited to only the 17% of patients undergoing resection(Speelman et al., 2015). FFPE sections from tumour excision biopsy are by far the most commonly used material in routine diagnostic laboratories due to difficulties in collection and storage of fresh or fresh-frozen samples. The formalin fixation process however, damages DNA through a number of mechanisms including fragmentation and cross-linking to proteins(Auerbach, Moutschen-Dahmen, & Moutschen, 1977). It is fortunate that the fragmented nucleic acids typically extracted from FFPE specimens are ideally suited, in length at least, to NGS platforms which are restricted to reading short length nucleic acids sequences also of around 200-225 base pairs(Shaw, Bullock, & Greenhalf, 2016). Despite the DNA damage accrued during the fixation process, studies have shown comparable sequencing quality with FFPE derived DNA compared to the gold standard of fresh or fresh-frozen samples(Shaw, Bullock and Greenhalf, 2016)(Spencer et al., 2013). A more significant problem of using FFPE for NGS is the tumour cellularity of the sample. Large scale sequencing studies using conventional approaches requires at least 80% tumour cellularity(Yadav et al., 2015). Dense desmoplastic stroma is a universal feature in PDAC which dilutes the mean tumour cellularity to between 38-44%(Mahadevan & Von Hoff, 2007). To some extent this can be overcome by coring out areas of high tumour cell content(Weng et al., 2010), either on the basis of gross histology(Wagle et al., 2012) or using histological guided laser capture

microscopy(Shen et al., 2014), but this is operator dependant and adds time to the workflow which may threaten clinical utility. These difficulties are illustrated in the Individualised Molecular Pancreatic Cancer Therapy (IMPaCT) study, the only trial to date investigating personalised therapy in PDAC(Chantrill et al., 2015). Low frequency sub-clones in the PDAC primary have an apparent tendency to be enriched in metastatic lesions(Campbell et al., 2010; Yachida et al., 2011), this perhaps reflects the greater metastatic potential of relatively slow growing cancer stem cells(Nies et al., 2014) but more importantly questions the rationale of directing therapy according to primary tumour sequencing in PDAC. Most concerning, and the potential death knell for use of solid tumour as a biospecimen in PDAC, is the demonstration that branched evolution produces driver mutations which are the most appropriate targets of therapy (Burrell et al.,, 2013; Maley et al., 2012b; Gupta et al., 2006)after cancer dissemination and therefore cannot be identified in the primary(Campbell et al., 2010). Despite these concerns sequencing of the primary tumour from FFPE is a useful exercise. It is not yet established that genomic information from CTCs can inform personalised therapy to improve outcomes in PDAC. Until this time, simultaneous sequencing of CTCs and FFPE primary tumour will be a useful correlation to direct future research efforts. Therefore, in this thesis, NGS of some primary tumour FFPE and CTCs is performed.

1.5.3. Circulating tumour cells

Given the limitations of solid tumour biospecimens, the use of blood as a biospecimen is highly attractive. Circulating tumour cells (CTCs) are cells shed from the primary tumour and found circulating in the vasculature, a sub-population of which may be capable of seeding distant metastasis(Gupta & Massagué, 2006). The dominant method for identification is the CellSearch™ system (Veridex) and because of this CTCs have come to be defined as cells isolated from blood with an intact nucleus, which stain positive for cytokeratin, epithelial cell adhesion molecule (EpCAM) and are negative for CD45(Andreopoulou et al., 2012). It has been over 10 years since CellSearch™, the only FDA approved CTC technology, first demonstrated the prognostic significance of CTC enumeration in metastatic breast cancer(Cristofanilli et al., 2004) and this has since been confirmed for lung(Hou et al., 2009), prostate(Stott et al., 2010) and colorectal cancer(Cohen et al., 2009). Studies have yet to demonstrate that CellSearch™ can improve survival by guiding treatment decision making. The phase III SWOG S0500 trial failed to improve survival using enumeration as a method of determining

chemotherapy efficacy in advanced breast cancer(Smerage et al., 2014). The DETECT study currently underway aims to determine whether treatment intervention guided by the HER2 status of CTCs in HER negative metastatic breast cancer patients (determined by primary tumour assessment) is superior to physician assessment(Schramm et al., 2015).

CellSearch™ relies on the expression of EpCAM and detects far fewer CTCs in PDAC than other cancers(Allard et al., 2005). EpCAM+ve CTCs are only found in 5% (4/75) of locally advanced PDAC (Bidard et al., 2013), 40% (21/53)(Khoja et al., 2012) and 42% (11/26)(Khoja et al., 2012) in mixed early and advanced cohorts and 48% (23/48) in advanced PDAC (Dotan et al., 2016). With its propensity for early haematogenous metastasis CTCs should be more abundant and more easily identified in PDAC than in other cancer types. Unfortunately, the converse is true; indeed, there is an inverse relationship between five-year survival in different types of cancer and EpCAM-based CTC recovery rates Figure 1-3

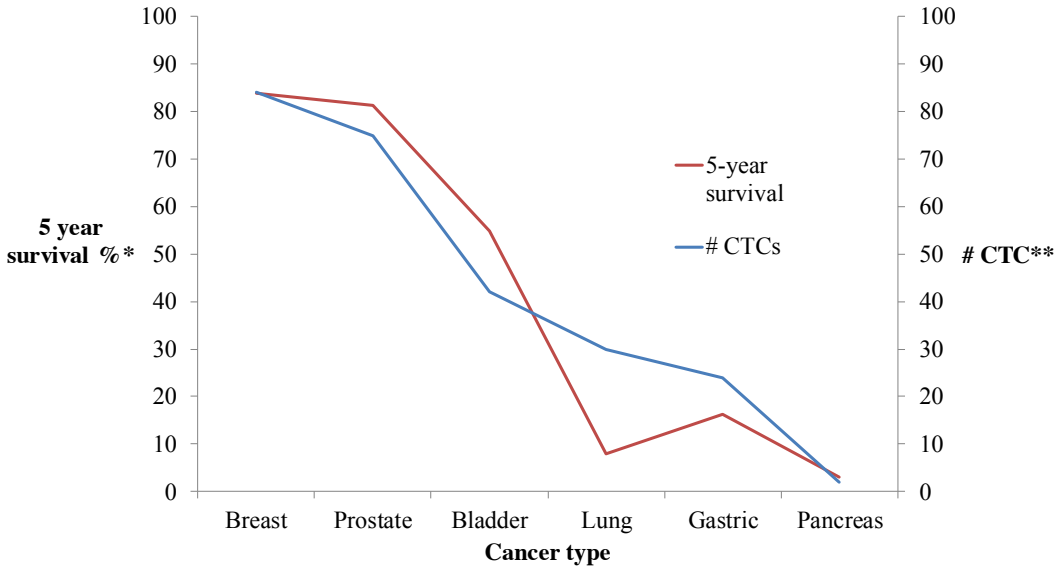


Figure 1-3 Correlation between mean number of EpCAM +ve CTCs and 5-year survival of cancer type. *(Quaresma, Coleman, & Rachet, 2015) based on 2005-2006 data, **(Allard et al., 2005) mean number of CTCs in metastatic cancer using CellSearch system

1.6 Epithelial mesenchymal transition (EMT)

The paradox described in the last section may be explained by epithelial mesenchymal transition (EMT): a process that makes tumour cells both prone to metastasise and evade EpCAM dependent detection. EMT plays an essential role in physiological processes such as embryology and tissue repair but is also implicated in the rapid formation of primary tumours(Mani et al., 2008)(Castellanos, Merchant, & Nagathihalli, 2013) metastasis(von Burstin et al., 2009), acquisition of therapeutic resistance(Arumugam et al., 2009) and poor survival(Yamada et al., 2013) associated with PDAC. In this process cells shed their epithelial antigens including EpCAM and cytokeratin (CK) and acquire mesenchymal markers such as COL5A2, EGFR, MSN, PDGFRB and Twist(Gorges et al., 2012). In mice the shedding of EpCAM from cancer cell lines injected intravenously occurs within 4 hours(Gorges et al., 2012) perhaps by cleavage of the extracellular EpCAM domain (the intracellular domain translocates to the nuclease and drives cell proliferation) upon endothelial cell contact(Denzel et al., 2009). A degree of phenotypic plasticity has been observed whereby cells may transition between epithelial and mesenchymal state with CTCs existing in both forms(Armstrong et al., 2011) with a purely mesenchymal phenotype predominating in the metastatic stages(Wu et al., 2015). This transient nature suggests that a reversible epigenetic mechanism rather than permanent genetic change is at play(Raimondi, Nicolazzo, Gradilone, Molecolare, & Università, 2015).

The most widely used CTC enrichment systems including CellSearch™, Adna Test(Demel et al., 2004), Magnetic Activated Cell Sorting System (MACS®)(Griwatz, Brandt, Assmann, & Zänker, 1995) and microfluidic technologies(Nagrath et al., 2011) all require cell surface expression of EpCAM for CTC capture and are therefore largely redundant in PDAC. Gorges et al, have advocated the use of new mesenchymal cell surface markers currently being pursued by Adna(Kasimir-Bauer, Hoffmann, Wallwiener, Kimmig, & Fehm, 2012), CellSearch™(Kasimir-Bauer et al., 2012) and CanPatrol CTC(Wu et al., 2015). However, positive selection with surface markers requires *a priori* identification rendering it vulnerable to missing CTC sub-populations without said cell surface markers(Satelli et al., 2014).

The only approach to overcome this is to adopt a method of CTC selection which avoids the fickle cell-surface markers completely. The genomic signature of the cancer cells are now well characterised in PDAC(Biankin et al., 2012; Campbell et al., 2010; Waddell et al., 2015) and can be used to both to distinguish CTCs from wild type cells

and potentially to guide therapy on the basis of actionable mutations. The main barrier to this approach is dilution of CTC genomic signal with wild type DNA from the vast excesses of nucleated white blood cells in the vasculature and a negative selection step, a means of removing white blood cells from a blood sample, is required.

1.7 CTC Enrichment

1.7.1. Negative selection

Whole blood is composed of red blood cells (RBC), white blood cells (WBC) and platelets suspended in plasma. Neither RBC nor platelets contain a nuclei or DNA and so will not contribute to the dilution of mutant DNA signal from CTCs. In health, there are between 4-11 million nuclei containing white blood cells per mL of whole blood(McKenzie SB, 1996) which vastly outnumber the 1-10 CTCs per mL which are typically quantified by EpCAM based technologies in cancer patients(Allard et al., 2005). The aim of a negative depletion is to reduce the WBC:CTC ratio from 6 logarithms, down to a level at which the CTC mutant signal can be detected; this will depend on the sensitivity of the detection instrument. The cluster of differentiation (CD) classification defines cells according to the molecules on their surface. CD45 is a receptor-linked protein tyrosine phosphate that is expressed, albeit in different isoforms on all nucleated WBC. CD45 has an extracellular domain of 400-550 amino acids(Altin & Sloan, 1997). The principle of negative depletion using magnetic antibodies involves the covalent coupling of anti-CD45 antibodies to streptavidin-coated magnetic beads, followed by mixing with whole blood to bind CD45 expressing WBC and finally extraction of WBC-bead complex from the sample with a magnetic field.
Rosette Sep^{TM} is an example of a non-magnetic negative depletion CTC enrichment which uses tetrameric antibody complexes against a variety of CD markers present on WBC at one end and glycophorin A on RBC at the other. The increased density of the bound cell complex (rosette) permits separation when centrifuged over a density gradient. The rossettes sediment with the RBC and the CTCs are negatively enriched in the CD45 negative mononuclear layer(Naume et al., 2004).

1.7.2. Selection by physical characteristics of CTCs

Other methods to enrich CTCs independent of EpCAM markers are by exploiting differences in size between WBC and CTCs. Solid tumour cell lines are generally

between 11.7µm and 23.8µm(Harouaka, Nisic, & Zheng, 2013) and even larger epithelial CTCs ranging from 29.8 to 33.9µm have been observed in metastatic breast cancer(Meng et al., 2004). WBC are generally smaller between 6.2-9.4µm(Harouaka et al., 2013). The CellSieve™ CTC isolation kit, (Creatv MicroTech, Canada) uses a 7µm filter pores in a micro filter created by the photolithographic fabrication method (D. Adams et al., 2015) trapping the CTCs. Caution should be exercised as CTCs from clinical samples may be smaller than those of cancer cell lines and the heterogeneity of cell surface markers is also likely to extend to cell size. To date this method has only been assessed with respect to cell line spiking experiments(D. L. Adams et al., 2014).
CTCs are believed to derive from two sources, firstly, via passive shedding directly from the primary tumour directly into the blood stream. These are likely to be large epithelial CTCs expressing EpCAM and fairly easily captured with size-based methods described above due to their greater diameter compared to WBCs. However their clinical significance is doubtful, not least because they will be too large to traverse the capillary bed (typically around 10 micrometers in diameter) and seed distant metastasis(Brabletz, Kalluri, Nieto, & Weinberg, 2018)(Nagrath, Jack, Sahai, & Simeone, 2016). Secondly, CTCs may derive by acquiring specific characteristics via the process of epithelial-mesenchymal transition (EMT)(Brabletz et al., 2018) discussed further in section 1.6. There is now mounting evidence that tumour cells acquiring the mesenchymal phenotype is an essential step in tumourigenic process(Brabletz et al., 2018). The smaller size of mesenchymal CTCs may enable them to pass through the capillary bed to seed distant metastasis(Brabletz et al., 2018) yet will escape capture by size based CTC enrichment techniques. Studies using sized-based CTC enrichment methods should therefore be interpreted with caution(Shibue & Weinberg, 2017).
The OncoQuik® system exploits the greater buoyancy of CTCs compared to other blood components by centrifugation over an optimized liquid separation medium. In validation studies with spiked cell lines it performed 100 times better at enriching spiked cells than standard Ficoll density gradient centrifugation(Gertler et al., 2003b). This method was outperformed by CellSearch™ in a small head to head study with low methodology quality(Balic et al., 2005) but CTC detection using this method was linked to disease progression in advanced breast cancer(Muller et al., 2005).

1.8 Biospecimen preparation

Regardless of the technical challenges of biospecimen optimisation, for clinical utility the institutional workflow must be conducive. Where batch processing of samples is a necessity to make the workflow viable, cryopreservation of fresh samples is needed. Cryopreservation is a reliable and convenient alternative to the use of fresh whole blood allowing optimal viability and functionality of cells and may be valuable when logistics require batch processing of samples(Ramachandran et al., 2012). There is a general consensus that gradual freezing leads to optimal viability of cells by minimising the formation of ice crystals both within and outside cells(Birkeland, 1976). There is less literature comparing the thawing conditions of cryopreserved cells with published protocols diverging after samples have been thawed in a 37°C water bath. The wash medium is added once the last remaining ice crystals are visible in some methods, and once thawing is completed in others. There is also wide variation in the temperature of washing medium, number of washes and speed at which washing medium is added. The optimum variables have been described in a systematic review of published protocols (Ramachandran et al., 2012).

1.9 Approaches to sequencing

1.9.1. Benchtop next generation sequencing

Identifying the cancer genome amongst the vast excesses of wild type (WBC) genomes is now feasible due to recent advances in sequencing technology. Massively parallel next generation sequencing (NGS) sequences millions of DNA templates simultaneously in a fraction of the time taken for Sanger sequencing. The two main NGS benchtop platforms, Ilumina and Ion Torrent™, both sequence DNA by recording the addition of nucleotides during DNA synthesis. Ion Torrent™ differs by monitoring pH rather than fluorescence during this process. A broad overview of the Ion Torrent™ sequencing platform used in this work is given here. Firstly, a specific known single stranded DNA sequence is attached many times to millions of beads, or Ion Sphere™ Particles (ISP). A complementary sequence is then used as an adapter and ligated onto the single stranded fragments in the library. They are then incorporated such that each ISP has its own DNA fragment. Clonal amplification is then performed by emulsion Polymerase chain reaction (PCR) for each fragment within the micro droplets, resulting in each ISP being covered by millions of copies of the same fragment. The ISPs are then

loaded onto the ion 316™ chip which contains 3 million microwells, ideally, one ISP is loaded per well. Data from polyclonal (more than one ISP per well) readings is discarded and empty wells produce no data. During the sequencing process, the chip, and therefore, each well is flooded with one of the four DNA nucleotides. If the nucleotide is complimentary to the next base in the fragment, the base is incorporated into the sequence by DNA polymerase, releasing a hydrogen ion in the process. When the base is not complimentary, no incorporation occurs and no hydrogen ion is released. If two (or x) bases are adjacent, twice (or x-fold) the number of hydrogen ions are released. Beneath each well is a metal oxide sensing layer overlying a sensing plate and floating metal gate that records the change in pH resulting from the hydrogen ion release and transmits the electrical signal to the semiconductor. This process occurs in parallel for each of the 3 million wells (assuming 100% loading) with a different nucleotide washing over the chip every 15 seconds. The error rates in this process are approximately 1% but maybe higher in high homopolymer (multiple consecutive nucleotides) DNA regions where it cannot distinguish the voltage of x base incorporations from x-1 bases incorporations.

1.9.2. Limiting dilution

When attempting to sequence low frequency variants such as seen in negatively depleted enriched CTCs, distinguishing low frequency variant signal from the inherent PCR and sequencing error described above becomes problematic. Limiting dilution is a method designed to overcome this problem and is used throughout this work, the concepts of which are described here. Suppose an enriched sample contains 1 CTC for every 100 WBC, the resulting DNA mixture would give a genuine variant frequency of 1%. If sequenced at 100% concentration, the 1% variant frequency would be indistinguishable from artefactual variants also at around 1% frequency. The limiting dilution method first dilutes the sample down to 10 genomes per µL and then aliquots 1µL into 10 separate wells. There are now 100 genomes, given a 1% genuine variant frequency, by chance there should be one genuine variant genome in one of the wells. All 10 wells are then sequenced. Artefactual variants will be expected across all 10 wells at the 1% frequency. Nine of the ten wells with only wild type genomes will show 0% genuine variant frequency; the one well containing the variant genome amongst 9 other wild type genomes will yield a variant frequency of 10%, referred to as the

'Jackpot' effect. Thus, by setting a variant threshold at 10% or above a genuine variant at 1% may be distinguished from artefacts. This example may be adapted to use different number of wells, dilutions and variant thresholds as required.

1.9.3. NGS structure

The sequencing capacity of a chip is finite, for example the ion 316™ chip has a 3 million microwells and each one of these can theoretically sequence a ~200bp DNA fragment; a 'read' (1/3 redundancy is typical). Whether one DNA fragment is sequenced 3 million times, or 3 million different DNA fragments are sequenced once can be tailored according to the clinical/scientific need. Other chips available for use on the Ion Personal Genome Machine (PGM)™ System include the Ion 318™ chip with 5.5 million wells and Ion 314™ chip with 0.5 million wells. A larger chip, Ion PI™ Chip has 80 million wells but requires the Ion Proton™ sequencing system. The three main categories of NGS sequencing are whole genome sequencing (WGS), whole exome sequencing (WES), and targeted sequencing.

WGS and WES permit a hypothesis free approach to somatic variant identification and have given valuable exploratory information for PDAC(Biankin et al., 2012; Nones et al., 2014; Waddell et al., 2015) and many other cancers. The vast amount of data dictates that a shallow depth of <100 reads is achieved and therefore demands a tumour cellularity of >70% which is challenging enough for primary PDAC but will not be possible with negatively enriched CTCs.

Targeted sequencing ranges from a single DNA fragment covering a SNP of choice to a comprehensive gene panel covering the exons from hundreds of genes. A narrower focus will permit a greater depth and therefore greater sensitivity. Targeting a gene that is mutated late in PanIN to PDAC progression has the advantage of improving specificity for cancer in analysis of circulating cells. *TP53* is ideal for PDAC as it occurs late and occurs in up to 70% of PDAC (R. H. Hruban, Goggins, Parsons, & Kern, 2000). The vast majority of mutations are found in exons 5-8(Rivlin, Brosh, Oren, & Rotter, 2011) allowing an even more focused approach and therefore much greater sequencing depth. As a frequent driver mutation *TP53* will more than likely maintain its presence through clonal sweeps as a trunk mutation(Greaves & Maley, 2012a), but despite its frequent presence in the primary, it is susceptible to being replaced as the dominant driver mutation by additional branch mutations(Campbell et al., 2010) thus limiting the single gene assay. Personalised therapy will only be possible if a selection

of genes which encompass the known heterogeneity of the tumour uncovered in WGS/WES studies can be made, preferably all the known actionable mutations. Oncology consortiums have recently developed custom gene panels using multiplex PCR combined with amplicon-based NGS from as little as 10ng of DNA derived from FFPE(Tops et al., 2015). The composition of the gene panels reflects both the frequency of mutated genes and oncogenes with potentially actionable mutations. Emphasis remains on validating diagnostic tests across multiple clinical laboratories(Dijkstra, Tops, Nagtegaal, van Krieken, & Ligtenberg, 2015; Tops et al., 2015), allowing in-house downstream bioinformatic analysis. For trials such as the National Cancer Institute (NCI) MATCH trial, larger custom panels including up to 200 genes are used on easily accessible platforms (such as the Ion Torrent PGM)(Redig & Janne, 2015). Though these panels are designed for sequencing of the primary tumour, application to alternative biospecimens such as enriched CTCs could be considered, and indeed is in this work, in PDAC. Expected tumour cellularity maybe as low as between 0.1-1% and depths of >500 will be required to differentiate from wild type DNA and inherent sequencing error. The downstream analysis for detecting the genomic signature of CTCs needs to be sensitive enough to overcome the diluting signal from wild type DNA regardless of enrichment type.

1.9.4. Variant assessment and nomenclature

In this thesis the human genome version 19 is used as a reference against which variants are called. However, this fails to account for individual variation of human genome. Variant filters are required to exclude synonymous likely benign variants and those, likely germline variants which occur frequently within the general population (minor allele frequency < 0.3). However, there still remains a significant number of non-synonymous variants which are not present to any degree in the general population which do not contribute to carcinogenesis. Secondly, the classification of variants as either driver or passenger(Stratton, Campbell, & Futreal, 2009) must be attempted as therapy directed on the basis of passenger mutations will be futile. Currently, only a qualitative assessment of the likelihood of pathogenicity is possible and is performed in three areas. Firstly using scores; the Sorting Intolerant from Tolerant (SIFT) score which uses sequence homology to predict whether such non-synonymous variants affect protein function(J. Zhang et al., 2014). Quality scoring tools such as the P value and its logarithmically related Phred quality score estimate the chance of the variant being due

to base calling error and hence reduce the number of variants included due to sequencing artefact. Secondly, determination of the change in protein effected by the variant. Ion Reporter™ software reports this automatically as either REF (sense), missense, or nonsense. Finally, cross referencing with the ever expanding variant databases such as COSMIC(Bamford et al., 2004) and dbSNP(Sherry, Ward, & Sirotkin, 1999).

1.10 Aims

1. Develop and optimise a technology to enrich and molecularly characterise CTCs in PDAC.

2. Use the technology developed to track the mutational profile of CTCs through a patient's treatment.

2 PATIENTS, MATERIALS AND METHODS

2.1 Pancreatic Juice as a potential source of biospecimen

2.1.1 Prophylaxis in ERCP and pancreatic juice collection

UK residents of the EUROPAC registry were invited to take part in the EUROPAC screening programme, a study approved by the relevant research ethics committees (LREC, AAGPM97199 [1998]; MREC, 07/H1211/96 [2007], Protocol Version 2 [2008]) and co-sponsored by the University of Liverpool and the Royal Liverpool and Broadgreen University National Health Service (NHS) Hospitals Trust. As part of the screening protocol, in addition to annual measurements of serum Ca 19-9 and regular imaging of the pancreas (computed tomography, magnetic resonance imaging and endoscopic ultrasound) patients were offered ERCP and collection of pancreatic juice for molecular analysis to stratify PDAC risk.

Between 6th January 1999 and 1st December 2013, 60 individuals at high risk of PDAC, 48 from FPC kindreds and 12 from HP kindreds underwent 80 ERCPs and collection of pancreatic juice for molecular analysis to stratify their PDAC risk further. This is in contrast to a later cohort of patients in the SSECRETIN study who underwent pancreatic juice collection using the duodenal aspiration method described later avoiding pancreatic duct cannulation.

ERCP was performed by consultant gastroenterologists at the Royal Liverpool and Broadgreen NHS Trust. 750 mg ciprofloxacin was given both before and after ERCP to minimise infective complications such as acute cholangitis(Raty S, Sand J, Pulkkinen M, 2001). Sedation was administered in the form of midazolam (1–5 mg) with hyoscine butyl-bromide (20–40 mg) as an antiperistaltic agent and either fentanyl (50–100 μg) or pethidine (25–50 mg). Selective cannulation of the pancreatic duct was confirmed by radiological screening without contrast followed by administration of 10 IU/kg secretin (Sanochemia, Germany) intravenously. After 2 minutes, pancreatic juice was collected by gentle aspiration from the pancreatic catheter. Pancreatic juice was analysed for *TP53* and *KRAS2* mutation and quantification of *CDKN2a* promoter methylation by Dr. Li Yan and described in detail in the groups paper(Yan et al., 2005).

Post-ERCP acute pancreatitis (PEP) was defined as a rise in serum amylase to at least 3 times (>450 IU/L) the upper limit of normal (150 IU/L) associated with epigastric pain within 48 hours of the ERCP procedure. Prophylaxis consisted of a 3-cm 5F self-expelling stent (Zimmon single pig-tail, no flap; Cook Medical) deployed using a 5F

introducer after aspiration of pancreatic juice and diclofenac administered per rectum within 30 minutes of the procedure.

Continuous data are presented as median and interquartile ranges (IQRs); categorical data are displayed as tables of counts and associated percentages. Associations of factors across patient groups were carried out using a 2-tailed Mann-Whitney U test for continuous data and Fisher's exact test for categorical variables. Risks are presented as odds ratio with associated 95% confidence intervals and are obtained from the parameters of univariate logistic regression models.

As EUROPAC research fellow from May 2013 to September 2014 my role in this study was to counsel and consent patients in the EUROPAC screening clinic. Following recruitment, I was responsible for the clinical supervision of patients undergoing ERCP, the injection of secretin, collection and processing (but not molecular analysis) of pancreatic juice and post procedure clinical care. I was involved in the collection of data for this study, but the primary analysis was performed by my EUROPAC research fellow predecessor Mr. James Nicholson. Within the EUROPAC study group it is accepted that EUROPAC research fellows work together and ensure their respective research interests overlap to a degree to ensure continuity of the project. I worked closely with both my EUROPAC processor Mr James Nicholson and EUROPAC successor Mrs Andrea Sheel to ensure continuity but have explicitly highlighted my role in the projects.

The samples collected within the SSECRETIN study are all analysed within a few days of collection. However, as stated previously, collection of pancreatic juice samples for the EUROPAC study began in 1999.These historic samples have always been collected and stored strictly within the GCLP guidelines and described in the study protocol. To ensure these samples have not degraded over the years to a degree which could affect their validity, historic samples underwent repeat testing at various subsequent time points using the same methodology and found to be concordant. This would suggest contemporary analysis and interpretation of these historic samples is valid.

2.1.2 Comparison of pancreatic juice supernatant vs pellet for biospecimen use.

The SSECRETIN study received ethical approval from the Haydock NRES Committee North West (REC 10/H1010/19). Patients undergoing an ERCP or endoscopic

ultrasound (EUS) procedure for suspected PDAC, chronic pancreatitis or biliary duct stones at the Royal Liverpool and Broadgreen NHS Trust were prospectively recruited between 2011 and 2015.

Following endoscopic intubation of the duodenum (no pancreatic duct cannulation) any fluid present was aspirated and discarded. 1 IU/kg secretin (Sanochemia, Germany) was then administered intravenously. After 2 minutes, pancreatic juice was collected by aspiration from the duodenum for either 10 minutes or until 10 mL of fluid was aspirated whichever came first. The ERCP/EUS then proceeded as per the initial indication. The pancreatic juice was then transferred to up to 10 Eppendorf 1.5 ml microcentrifuge tubes (Sigma-Aldrich, referred to from here on as Eppendorf tubes) and centrifuged at 300g for 10 minutes. The supernatant was transferred to Nunc cryotubes (Thermofisher, referred to from here on as Nunc tubes) and both supernatant and pellet were stored at -80˚C. My role in this study was patient identification, recruitment, administration of secretin and collection of pancreatic juice in a proportion of the patients. I processed the pancreatic juice and stored it in a -80^{o}C freezer. Molecular analysis was performed by Dr. Li Yan postdoctoral researcher and Miss Hollie Pufal undergraduate student.

The pellet was thawed and re-constituted with 200µL of PBS, the supernatant was thawed without dilution. 200µL of the each was inputted to the MagNA Pure Compact Instrument (cat. no. 03731146001, Roche, UK) to elute DNA according to manufacturer's instructions. Molecular analysis was carried out for *KRAS* by Amplification Refractory Mutation System (ARMS) PCR, and real time PCR measurement of *CDKN2a* promoter methylation, described in full in the group's previous publication(Yan et al., 2005) and described in brief below.

2.1.2.1 KRAS ARMS PCR

Real-time polymerase chain reaction (PCR) was performed using the LightCycler 480 with SYBR Green detection (Roche Diagnostics, Pinzberg, Germany). Because there is amplification of the wild-type sequence with mutation-specific primers, the threshold cycles with mutant-specific primers were plotted against threshold cycles using control primers that amplified both wild-type and mutant sequences. PCR of 100 blood samples allowed 98% confidence intervals to be produced on a linear regression curve for each mutation-specific primer. Samples were analysed in triplicate; if all 3 points were less

than the 98% confidence limit then the sample was classified as mutant(Yan et al., 2005).

2.1.2.2 CDK2Na promoter methylation

Pancreatic juice DNA was sodium-bisulfite modified. PCR amplification was performed again using the LightCycler 480. The methylation index was [M/(M + U)] x 100%, where M is the quantity of methylated *CDKN2a* promoter and U is the quantity of un-methylated *CDKN2a* promoter measured by real-time PCR.

2.1.3 Comparison of the molecular profile of whole pancreatic juice to the pellet and supernatant

Whilst the SSECRETIN study was underway and during my time as EUROPAC research fellow studies emerged suggesting that the use of whole duodenal juice was the optimal biospecimen to use in this context(Mitsuro Kanda, Knight, et al., 2013) A modification to the processing of collected pancreatic juice was made in July 2014 to allow comparison of whole juice with that of pellet and supernatant; a minor amendment which I submitted was accepted by the ethics committee (protocol version 2). After collection of the sample, 2mL of whole juice was placed into two Nunc tubes and stored at -80°C. The remainder to the juice was processed in the manner described previously i.e. transferred to up to 8 Eppendorf tubes and centrifuged at 300g for 10 minutes. The supernatant was transferred to Nunc tubes and both supernatant and pellet stored, along with the whole juice at -80°C.

2.2 A negative depletion approach to using blood as a biospecimen

2.2.1 Small volume spiking experiments

All depletions were performed by Dr. Nick Bryan at the Department of Clinical Engineering, Liverpool University. For peripheral blood mononuclear cell(PBMC) harvesting, 7mL of healthy volunteer whole blood was obtained with informed consent in a 7.5mL Ethylenediaminetetraacetic acid (EDTA) tube (cat. no. 7mL 01.1605.001, Sarstedt, UK) and diluted 1:2 in phosphate buffered saline (PBS) (Sigma-Aldrich, UK), layered over 15mL of lymphoprep™ (cat. no. 07811, Stem cell technologies, UK) and centrifuged at 400g for 30 minutes at 20 °C. The monocyte layer at the interphase was

carefully removed into a 50mL tube, topped up with PBS and centrifuged once more at 300g for 10 minutes at 20°C discarding the supernatant.

The PANC-1 cell line (cat. no. 87092802, Sigma-Aldrich, UK) was prepared and cell culture maintained by colleague Dr. Li Yan. The cells were cultured in RPMI-1640 Medium (cat. No. R8758, Sigma-Aldrich, UK) supplemented with Foetal Bovine Serum (FBS, cat.no. 12003C, Sigma-Aldrich, UK) and glutamate. At approximately 80% confluency the cells were trypsinised, washed and pelleted by spinning at 1000 x g for 10 minutes. Mycoplasma testing was not routinely performed. Genomic sequencing of the cell line was performed at various dilutions using NGS and confirmed the expected mutations and purity of the cell line and described in full in 2.4.1. All other cell lines including MIA-Pa-Ca-2 (cat. no. 85062806, Sigma-Aldrich, UK) and SUIT-2

Cell counting for both PANC-1 and PBMCs were performed using a Nucleocounter according to manufactures guidelines. Two spiked sample groups were prepared; 30,000 PANC-1 cells spiked into 200,000 PBMCs and 15,000 PANC-1 cells spiked into 30µL whole blood diluted 1:10 with PBS Table 2-1.

Anti-human CD45 antibodies were covalently coupled to streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Dynal, Oslo, Norway) according to manufacturer's guidelines by Dr. Nick Bryan. For red blood cell depletions, the protocol above was followed substituting a red blood cell antibody for the CD45 antibody.

The bead and antibody complex was then added to the samples and briefly vortexed. Following 20 minutes incubation at room temperature a magnet was passed slowly over the sample visibly removing the magnetic beads. The samples and various depletion types are described in Table 2-1.

Table 2-1. Experimental set-up of small volume spiked experiments.

No. of PANC-1 cells	Sample	Depletions
30,000	200,000 PBMCs	No depletion
		1 x CD45
		2 x CD45
		3 x CD45
15,000	30µLwhole blood (1:10 PBS)	No depletion
		1 x RBC & 1 x CD45
		2 x RBC & 1 x CD45
		3 x RBC
		3 x RBC & 1 x CD45

Following the CD45 depletions DNA elution was performed by placing 200µL of each sample directly into the MagNA Pure compact instrument (cat.no. 03731146001, Roche, UK) according to manufacturer's guidelines. Standard real time PCR was then performed on the eluted DNA using the LightCycler480 (cat. no. 5015278001, Roche, UK). For each 25µL reaction the following reagents were used: 2.5µL of Gold Buffer (cat. no. P2317, Sigma, UK), 1.56µL of $MgCl_2$ (cat. no. M8787, Sigma, UK), 0.52µL of dNTP mix (cat. no. dNTP100A, Sigma, UK), 0.4µL forward primer (wild type *KRAS*: 5'TGA CTG AAT ATA AAC TTG TGG TAG TTG GCG3'), 0.4µL common reverse primer (5'CTC ATG AAA ATG GTC AGA GAA ACC TTT ATC3'), 17.491µL of molecular grade DNA free water (cat. no. W4502-11, Sigma, UK) and 2µL of DNA. The PCR parameters were as follows: pre-incubation for 13 minutes at 95°C, DNA amplified for 40 cycles of 95°C for 50 seconds, 54°C for 45 seconds, 72°C for 1 minute, and finally 10 minutes at 72°C. Each sample was amplified in triplicate for wild type *KRAS*, triplicate for p.G12D (PANC-1) *KRAS* and triplicate for p.G12R *KRAS* as a negative control. Water controls were also used for each 96-well plate.

2.2.2 Antibody and bead ratio

Healthy volunteer whole blood was obtained in the same manner and prepared into aliquots of 30µL. The same anti-human CD45 antibodies were covalently coupled once again to streptavidin-coated magnetic beads but with 5 different antibody volume variables: 0µL, 10µL, 50µL, 100µL and 120µL and two different bead volume variables: 20µL and 40µL to make 10 separate bead-antibody complexes. CD45 depletion was otherwise performed in the same manner on 10 x 30µL aliquots of whole blood. For each sample, both the beads and the remaining supernatant were kept making 20 samples. The beads and supernatant samples were suspended in 200µL of PBS and placed into the MagNA pure system to elute DNA into 100µL. DNA was then quantified, each sample in triplicate using the Qubit 2.0 fluorimeter. Linear regression analysis was used to study the antibody volume relationship. Students paired t test was used to compare concentration of DNA according to bead volumes.

2.2.3 Ion Torrent™ sensitivity

The sensitivity of the Ion Personal Genome Machine® (PGM™) System was investigated by Mr. James Nicholson. PANC-1 DNA from the same source was made into experimental samples by 10-fold serial dilution with wild type DNA. The ratio of Wild type:PANC1 DNA of the samples was pure PANC-1, 900:100, 990:10 and 999:1. Each sample was then diluted to 10 genomes/µL and sequenced for *TP53* as described in full later. The frequency of the known PANC-1 variant p.R273H for each of the samples was compared with the proportion of PANC-1 DNA in the sample in each of the 10 barcodes.

2.2.4 Limiting dilution assessment

The effect of limiting dilution on variant frequency in the setting of detecting a CTC variant signal in whole blood was investigated separately. 1,000, 100 and 10 PANC-1 cells were spiked into 7 mL of healthy volunteer whole blood. A standard CD45 depletion (described earlier) was performed and DNA eluted. The control samples were sequenced, 10 barcodes each at a standard concentration of 5ng/µL. For the comparator, the same DNA was diluted to 10 genomes/µL (0.06ng/µL). Sequencing was performed

in the same way with 10 barcodes for each of the three samples. The frequency of the p.R273H variant was compared in each of the barcodes between standard and diluted samples. The means of the highest barcode frequencies were compared between groups using the paired t-test.

2.2.5 Large volume spiked CD45 depletions

7mL aliquots of healthy volunteer whole blood were spiked with 1,000, 100 and 10 PANC-1 cells and gently vortexed. A standard CD45 depletion was performed on each using CD4 antibody and the magnetic beads. Three controls were used: pure PANC-1 cells, 7mL blood with no spiked PANC-1 cells and 7mL blood with 1,000 spiked PANC-1 cells and depletion performed omitting the CD45 antibody component.

2.2.6 Large volume clinical CD45 depletions

All clinical blood samples were obtained by myself under the framework of the study: 'Comparison of techniques for detection of circulating tumour cells in peripheral blood 'with the Research Ethics Committee reference: 08/H1011/36, amendment 4, dated 11th June 2013. The samples were obtained with informed consent using the patient information sheet version 5, dated 11th June 2103 (appendix 1) and the consent form version 5, dated 9th August 2013 (appendix 2). Sodium Heparin tubes 7.5 mL (cat. no. 01.1613.100, Sarstedt, UK) were used for blood sample collection and all patients were current in-patients at the Royal Liverpool University Hospital at the time of recruitment. Standard CD45 depletions were performed by Dr. Nick Bryan within 4 hours of collection. To accommodate larger blood volumes the volume of beads and antibody were scaled up accordingly. Various experimental controls were used by omitting the antibody and/or beads.

2.2.7 *TP53* NGS library preparation and downstream analysis

DNA is initially eluted from the sample using the MagNA pure system. Where whole blood is used the sample is first diluted 1:1 with PBS to make a 400µL starting volume eluted to 200µL of DNA. DNA is then quantified using the Qubit 2.0 fluorimeter in triplicate with the mean calculated. The DNA is then diluted to 10 genomes/µL (0.06ng/µL) with nuclease free water (cat. No. AM9939, Ambion™, UK) and re-quantified to ensure appropriate dilution. Typically, a 700µL 10 genome/µL diluted

sample was prepared for downstream analysis. The nuclease free water volume:sample volume ratio for such a dilution was calculated by:

$$Water\ volume\ required\ for\ dilution = 700 - \left(\frac{700}{DNA\ conc/_{0.06}}\right)$$

then,

$$Sample\ volume = 700 - Water\ volume\ required\ for\ dilution$$

Primers for *TP53* NGS were designed and ordered by Mr. James Nicholson from Eurofins genomics, UK. Considered in the design of primers is optimal length of DNA fragment for NGS which is approximately 220 base pairs (bp). As 30 patients are multiplexed onto one chip, each fragment must have a unique barcode which can be recognised by the Ion Torrent™ software and each barcode is ligated onto the fragment via an adapter, typically 5-6 base pairs long. Exons 5, 7 and 8 of the *TP53* gene are all longer than 200bp and therefore had to amplified in overlapping fragments. As each fragment of DNA consists of two complimentary strands, primers for both strands were designed and labelled**,** according to convention as 'A' and 'P'.
The product length of each of the fragments (including primers) is shown in Table 2-2. Each of the 16 fragments Table 2-2 had 30 unique barcodes incorporated into the forward primer such that 480 distinct sequences were amplified.

Table 2-2. Product length of all TP53 fragments amplified by the designed primers

TP53 Primer	Product length
Exon 5 Fragment 1 A	217
Exon 5 Fragment 1 P	210
Exon 5 Fragment 2 A	227
Exon 5 Fragment 2 P	220
Exon 5 Fragment 3 A	191
Exon 5 Fragment 3 P	186
Exon 6 Fragment A	218
Exon 6 Fragment P	211
Exon 7 Fragment 1 A	196
Exon 7 Fragment 1 P	189
Exon 7 Fragment 2 A	215
Exon 7 Fragment 2 P	208
Exon 8 Fragment 1 A	200
Exon 8 Fragment 1 P	193
Exon 8 Fragment 2 A	221
Exon 8 Fragment 2 P	214

The PCR parameters were as follows: pre-incubation for 13 minutes at 95°C, DNA amplified for 60 cycles of 92°C for 15 seconds, 64°C (for exons 5 and 6) or 55°C (for exons 7 and 8) for 15 seconds, 72°C for 10 seconds, and finally 10 minutes at 70°C. The reagents used were 2.5µL of AmpliTaq Gold® with Buffer (cat. no. P2317, Sigma, UK), 1.56µL of $MgCl_2$ (cat. no. M8787, Sigma, UK), 0.52 µL of dNTP mix (cat. no. dNTP100A, Sigma, UK), 0.4µL of forward primer, 0.4µL of reverse primer, 17.49µL of molecular grade nuclease free water (cat. no. W4502-11, Sigma, UK) and 2µL of DNA. For every 96-well plate a negative control column was used substituting nuclease free water for DNA. Following amplification, 5µL of product (one of ten barcodes only) was combined with 2µL of loading dye (cat. no. R0631, Thermo Fisher, UK) and run on a 1.8% agarose gel at 400mA for 20 minutes with an appropriate ladder (cat. no.

11336045001, Roche, UK) to confirm amplification of product at an expected size and absence of water control well amplification to exclude DNA contamination Figure 2-1.

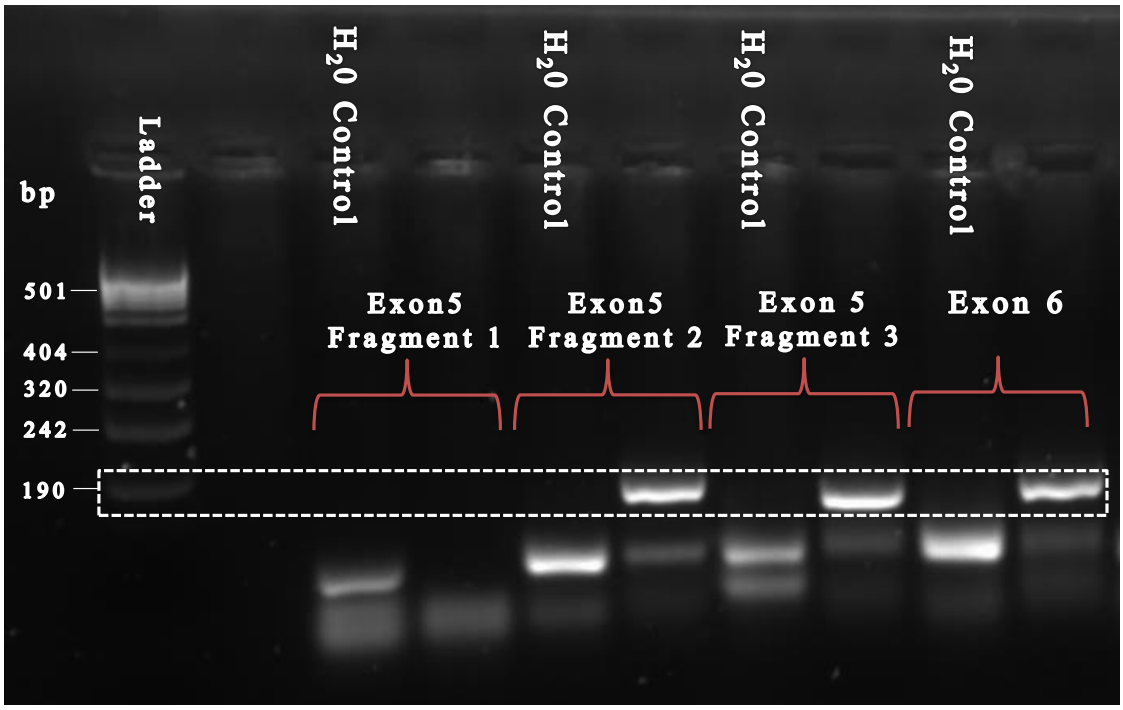


Figure 2-1. Agarose gel shows bands at expected size for amplification of exon fragments ~ 200bp (white dotted line) with absence of product in the H_20 control well. Exon 5 fragment 1 has not sufficiently amplified in this instance. More distal bands represent shorter products of unwanted primer-dimer products and are discarded

Purification was then performed on the remaining 20µL of product. All 10 barcodes and both A and P directions were pooled for each fragment making 8 separate pools of 200µL. 150µL of each pool (2 x 75µL) was loaded into a large 1.8% agarose gel combined with 10µL of loading dye at 400mA for 40 minutes with an appropriate ladder. The band of interest Figure 2-1 white dotted line, was then excised with a scalpel. DNA was extracted using the Agarose Gel DNA Extraction Kit (cat. no. 11 696 505 001, Roche, UK) according to manufacturer's guidelines. Each fragment was then quantified using the Qubit 2.0 fluorimeter. Appropriate volumes of each fragment were then combined according to their concentration to ensure an equal concentration of DNA for each fragment such that no fragment dominates the chip.

The sequencing was kindly performed by technician Miss Katie Bullock on the Ion Personal Genome Machine® (PGM™) System (cat. no 4462921, Life Technologies,

UK) Figure 2-2B according to manufacturer's guidelines and detailed in the Standard Operating Procedure (SOP) GCLPEQU022 - Use of the Ion Torrent™ (this and other SOPs are document controlled as part of Good Clinical Practice by the Cancer Research UK Liverpool Cancer Trials Unit). The following kits and reagents were used: Ion PGM™ Sequencing 200 v2 Kit (cat. no. 4482002, Life Technologies, UK) Figure 2-2A, Ion 316™ Chip v2 (cat. no. 4483188 Life Technologies, UK) C, Ion PGM Hi-Q OT2 Kit (cat. no. A27739, Life Technologies, UK) Figure 2-2D.

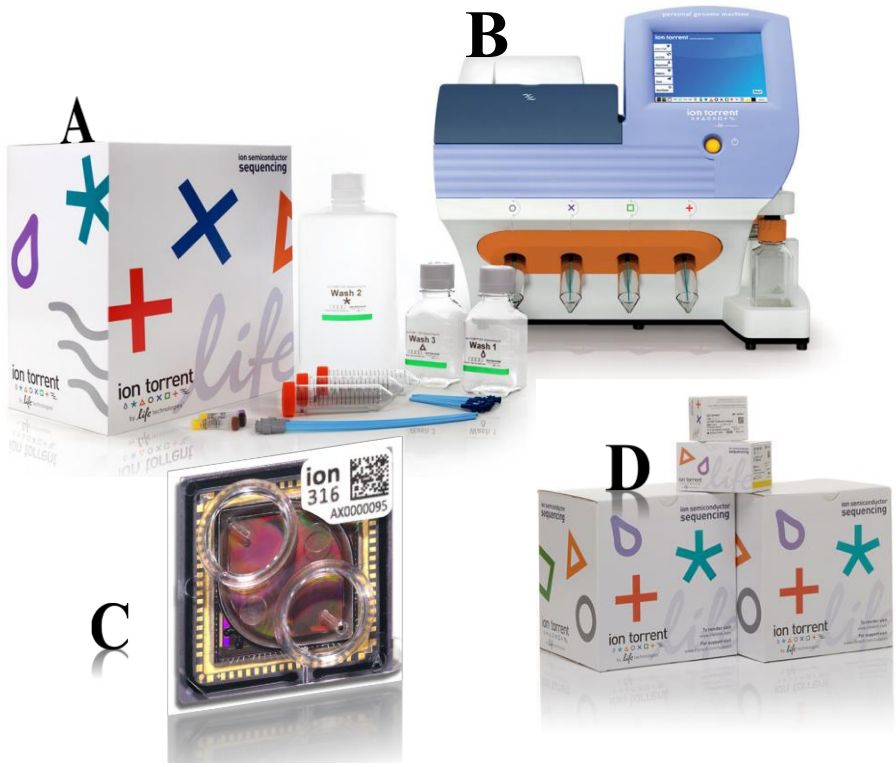


Figure 2-2. Ion Torrent™ equipment and reagents. A) Ion PGM™ Sequencing 200 v2 Kit, B) Ion Personal Genome Machine® (PGM™) System, C) Ion 316™ Chip v2, D) Ion PGM Hi-Q OT2 Kit

Sequencing runs were each downloaded individually to the Integrative Genomics Viewer (IGV), a visualization tool for genomic interrogation(Robinson et al., 2012). Each base pair for every barcode was individually reviewed, with variants occurring at greater than 10% considered significant where the limiting dilution parameters were set to 10genomes/µL and 10 replicates. Where the presence of a known variant is sought, such as repeat samples then a lower threshold is accepted. Variants were then cross referenced with the International Agency for Research on Cancer (IARC) *TP53*

database available at http://p53.iarc.fr/ which compiles all reported TP53 mutations in sporadic cancers and describes, the expected tumour phenotype and functional and structural impact of variants.

2.3 Assessment of three commercial CTC enrichment methods

2.3.1 Patient recruitment

All patients were recruited under the auspices of the study: 'Comparison of techniques for detection of circulating tumour cells in peripheral blood' with the Research Ethics Committee reference: 08/H1011/36, amendment 4, dated 11th June 2013 and informed consent outlined in chapter 2. Sodium Heparin tubes 7.5 mL (cat. no. 01.1613.100, Sarstedt, UK) were used for blood sample collection and all patients were current surgical in-patients at the Royal Liverpool NHS Trust at the time of recruitment.

2.3.2 CellSieve™

The CellSieve™ CTC isolation kit(D. L. Adams et al., 2014), (Creatv MicroTech, Canada) was performed by post-doctoral researcher Dr. Nick Bryan according to manufacturer's guidelines. In brief, 7.5 mL of fresh whole blood was used for each sample and processed within 4 hours of collection. The filter membrane was assembled into the filter holder shiny side up as shown in Figure 2-3A. 5mL of Foetal Bovine Serum (FBS, cat. no. 12003C, Sigma-Aldrich, UK) was drawn into the 50mL 'waste syringe' and attached to the syringe pump Figure 2-3B. The filter holder was then secured with a twisting motion on top of the waste syringe. The plunger was removed from a 30mL 'input' syringe and securely attached to the filter apparatus above Figure 2-3B, again, with a twisting motion. The blood sample was pipetted into the 30mL input syringe on top. The sample was drawn through the filter by the syringe pump set at 5mL/minute with a target volume of 16 mL, and force limit of 20%. After 1-2 minutes an additional 5mL of PBS was added to the input syringe and allowed to pass through at the same rate. This process was repeated twice before stopping the pump.

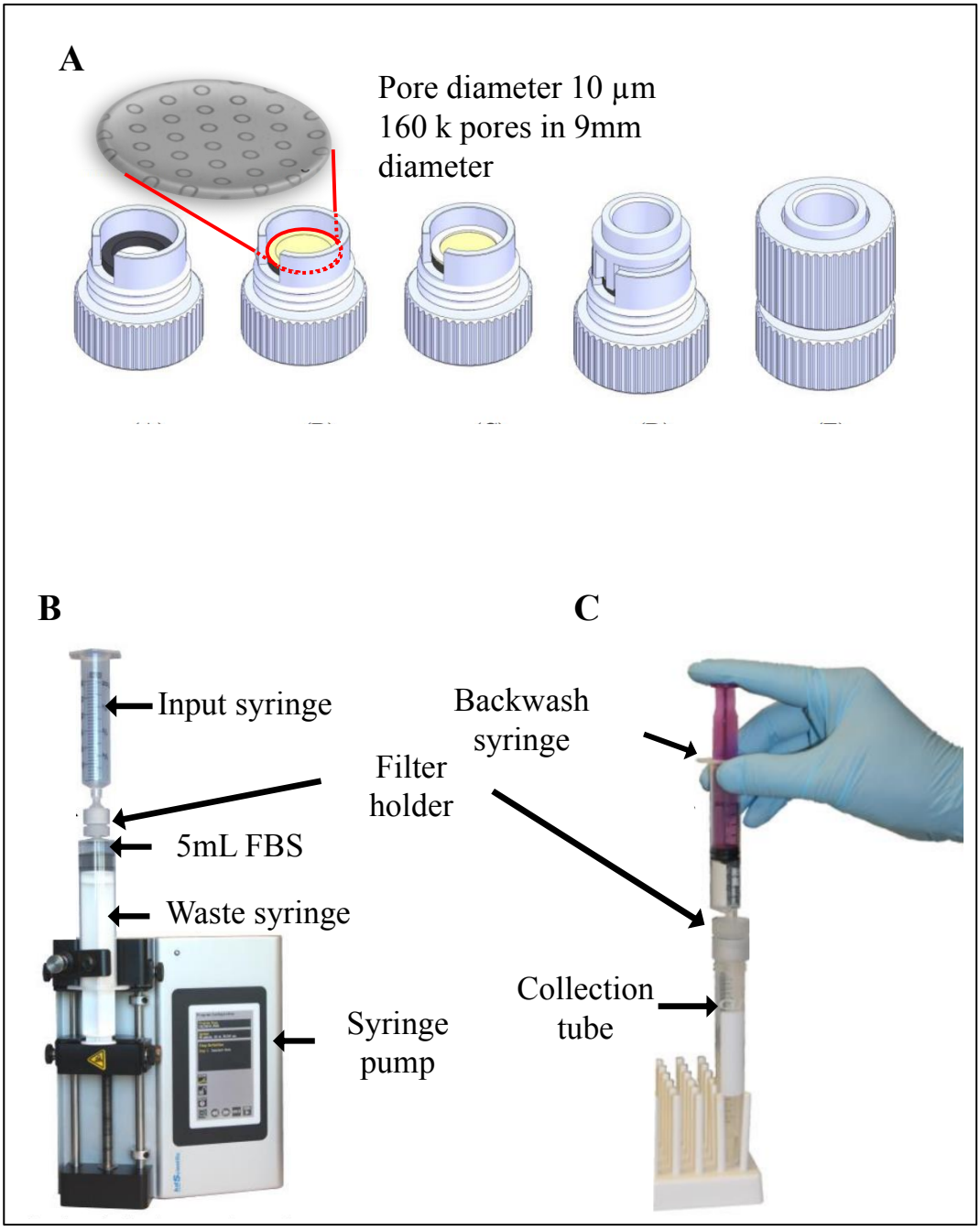


Figure 2-3. Arrangement of apparatus for the CellSieve™ micro filter. A) The filter (red circle) was mounted into the filter holder as shown. B) The apparatus set-up is shown with the input and waste syringe separated by the filter holder. C) The filter holder was inverted over a collection tube and the back wash syringe containing 5mL of PBS was used to flush any remaining enriched cells into the collection tube.

The filter holder was then removed and a 10mL 'backwash syringe' containing 5mL of PBS was attached to the side formerly occupied by the waste syringe Figure 2-3C. The filter holder was inverted over a collection tube and the syringe depressed at approximately 5mL/15 seconds flushing the enriched cells back from the filter and washing into the collection tube. The enriched sample was centrifuged at 1,200g for 10 minutes, the supernatant discarded and the pellet reconstituted in 200µL of PBS. The < 10µm non-enriched fraction beneath was collected from the waste syringe. DNA elution was performed by placing 200µL of each fraction directly into the MagNA pure system. The concentration of eluted DNA was then measured using the Qubit Fluorimeter 2.0 and diluted to 10 genomes/µL with nuclease free water (cat. No. AM9939, Ambion™, UK) and 5 sequencing replicates were performed for each sample for *TP53* analysis. *KRAS* analysis was performed using ARMS PCR, described previously, with undiluted DNA by Dr. Li Yan.

2.3.2 RosetteSep™

Three different RosetteSep™ enrichment cocktails were used: RosetteSep™ CTC Enrichment Cocktail Containing Anti-CD 36 (cat. No. 15167, Stemcell technologies, UK), Anti-CD 45 (cat. No. 15122, Stemcell technologies, UK) and Anti-CD 56 (cat. No. 15177, Stemcell technologies, UK) each containing various tetrameric antibody complexes recognizing glycophorin on red blood cells and a selection of antigens present on white blood cells shown in Table 2-3.

Table 2-3. Antibody complexes contained within each of the three varieties of Rosette Sep™ CTC enrichment cocktails.

Rosette Sep™ CTC Enrichment Cocktail	CD antigen										
	45	66b	38	16	19	36	2	3	14	56	61
Anti-CD 45	✓	✓									
Anti-CD 36	✓	✓	✓	✓	✓	✓	✓				
Anti-CD 56	✓	✓	✓	✓	✓			✓	✓	✓	✓

The method for each kit is identical except for the addition of the different CTC enrichment cocktail. 350µL (50µL/mL blood) of the cocktail was added to the 7mL

blood aliquot Figure 2-4A, briefly vortexed and incubated at room temperature for 20 minutes. The sample was then combined with an equal volume (7mL) of 2% foetal calf serum (FCS, Sigma-Aldrich, UK)/ phosphate buffered saline (PBS, Sigma-Aldrich, UK) and layered carefully on top of 15mL of lymphoprep^{TM} (cat. no. 07811, Stem cell technologies, UK) in a SepMate^{TM} blood separation tube (cat. No. 15450, Stem cell technologies, UK) Figure 2-4C. The tube was centrifuged at 1200g for 20 minutes with a brake setting of 5. The enriched fraction, above the SepMate^{TM} tube filter Figure 2-4D was removed and placed into a clean 15mL falcon tube and further washed with 2% FCS/PBS. The washed cells were then centrifuged at 300g for 8 minutes and the supernatant was discarded. The pellet of enriched cells was reconstituted in 200 µL of PBS and used directly to elute DNA using the MagNA pure system described previously. The pellet at the bottom of the falcon tube Figure 2-4D, representing the non-enriched fraction was mixed 1:1 with PBS. DNA was eleuted in the same way using 200µL. DNA was quantified using the Qubit 2.0 fluorimeter and subsequently diluted to 20 genomes/µL.

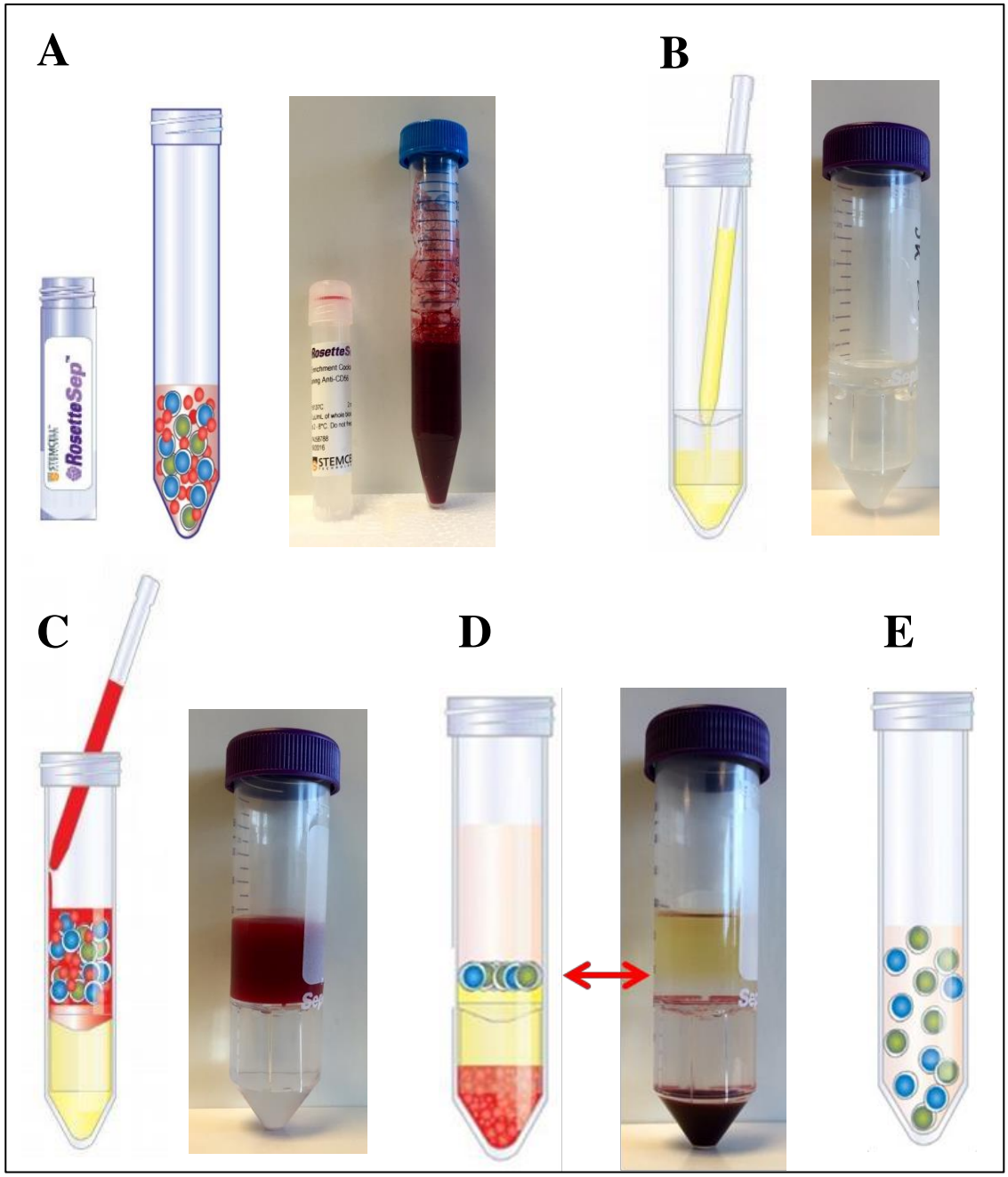


Figure 2-4. Method for RosetteSep™ CTC enrichment. A) RosetteSep™ CTC enrichment cocktail incubated with whole blood. B) Density gradient (Lymphoprep) added to SepMate tube. C) Whole blood layered in the upper compartment of the SepMate tube on top of Lymphoprep. D) Enriched visible layer (red arrow) after centrifugation. E) Enriched CTC fraction removed to clean tube.

2.3.3 OncoQuik®

OncoQuik® samples were processed according to manufacturer's guidelines. In brief, OncoQuik® tubes (cat. no. 227250, Greiner bio-one, Germany) were chilled on ice

Figure 2-5A. Whole blood was layered gently into the upper compartment above the porous filter Figure 2-5B. The tube was centrifuged at 1398g at 4^{o}C for 20 minutes with acceleration = 3 and no brake. The enriched portion (at the interphase between the plasma above and separation medium below) arrowed Figure 2-5C was then transferred to a clean falcon tube

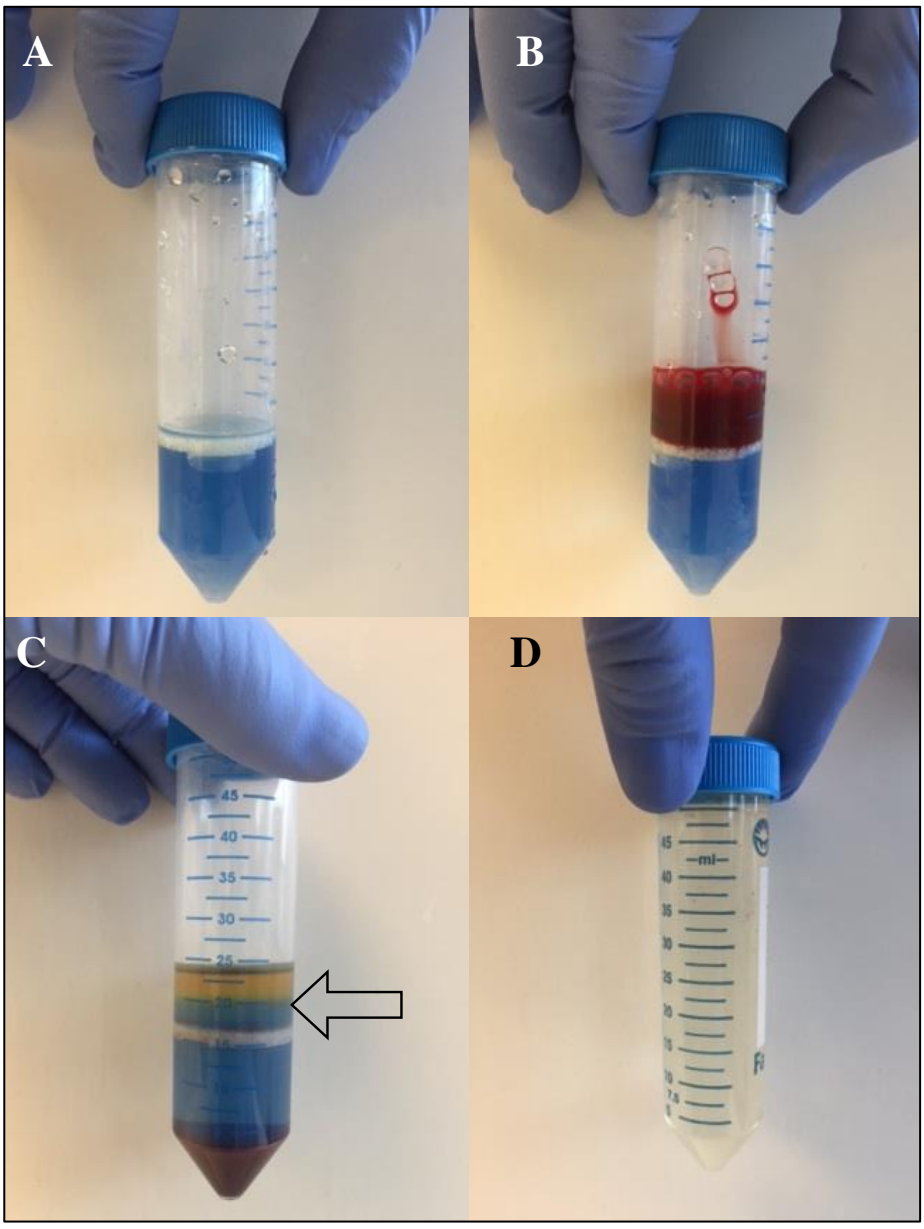


Figure 2-5. OncoQuik® enrichment protocol. A) OncoQuik® tube after chilled on ice. B) Whole blood layered in top tube compartment. C) After centrifugation, enriched CTC layer seen at interphase between plasma and density medium. D) Enriched cells with additional washing buffer up to 50mL.

and made up to 50mL with washing buffer, 0.5% (w/v) bovine serum albumin (BSA, Sigma-Aldrich, UK) and phosphate buffered saline (PBS, Sigma-Aldrich, UK) Figure 2-5D and pelleted by centrifuging at 200g for 10 minutes. The supernatant was discarded and the pellet reconstituted with 200µL PBS. The pellet at the bottom of the OncoQuik® tube, representing the non-enriched portion Figure 2-5C was mixed 1:1 with PBS. 200µL of both enriched and non-enriched fractions were used directly to elute DNA using the MagNA pure system.

2.3.4 Fractionation experiment

15 mL of whole blood was sampled pre-operatively from patient 10 on the morning of surgery for borderline resectable PDAC. This was divided into two equal aliquots of 7.5mL. The first aliquot underwent standard enrichment with OncoQuik® Figure 2-5. The second aliquot underwent a fractionation process described here: The 7.5 mL of whole blood was carefully layered over 15mL Lymphoprep™ Figure 2-6A and centrifuged at 1200g for 20 minutes with the brake on Figure 2-6B. Following centrifugation, 1 mL aliquots were aspirated from the very top of the tube and placed in a separate Eppendorf tube. This was repeated until 21 separate samples were obtained Figure 2-6C.

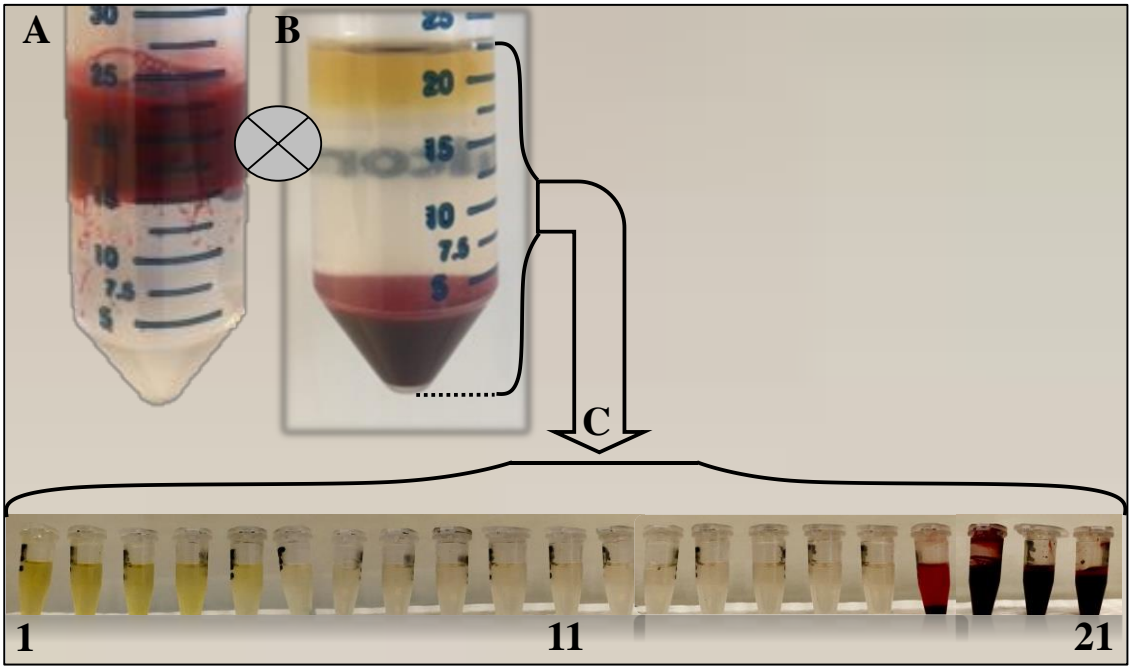


Figure 2-6. First fractionation process. A) Whole blood layered over Lymphoprep™. B) 1mL aliquots taken from top layer down into 21 separate Eppendorfs C.

The process of OncoQuik® enrichment and fractionation was repeated for the same patient on the first post-operative day. In both instances, the enriched OncoQuik® pellet was re-suspended in 400µL of PBS before DNA elution with MagNa pure system. It was then diluted to 10 genomes/µL and 2µL used in each sequencing replicate. 9 replicates were used for each of the OncoQuik® samples (permitting a total of 30 barcodes to be used including the 21 aliquots). DNA concentration was measured for each of the 21 fractionated samples using the Qubit 2.0 fluorimeter. 400µL of each of the fractionations was used without dilution for DNA elution using the MagNa pure system. A total of 10ng of DNA was used for each of the 21 fractions with no replicates (one barcode each).

2.3.5 OncoQuik® and RosetteSep™ comparison

19mL of whole blood was divided into 3 aliquots and underwent the following enrichments:

1. 7mL RosetteSep™ CTC enrichment using cocktail Anti-CD56
2. 7mL RosetteSep™ CTC enrichment using cocktail Anti-CD36
3. 5mL OncoQuik® enrichment

The enriched cells were sequenced in the manner described previously, they were diluted to 10 genomes/µL with 5 replicates per sample making 30 barcodes in total. Primers for the *TP53* variants identified from sequencing were designed with the aim to confirm their presence using mutation specific PCR.

2.3.6 TP53 primer design

Software from the primer design website available at www.primer3plus.com was used. The *TP53* sequence including the variant site was copied and pasted into the template sequence box. The forward primer was selected by identifying the variant nucleotide of interest and including the preceding 19 base pairs. An example for forward primer for the variant c.814G>A, p.V272M is shown in Figure 2-7.

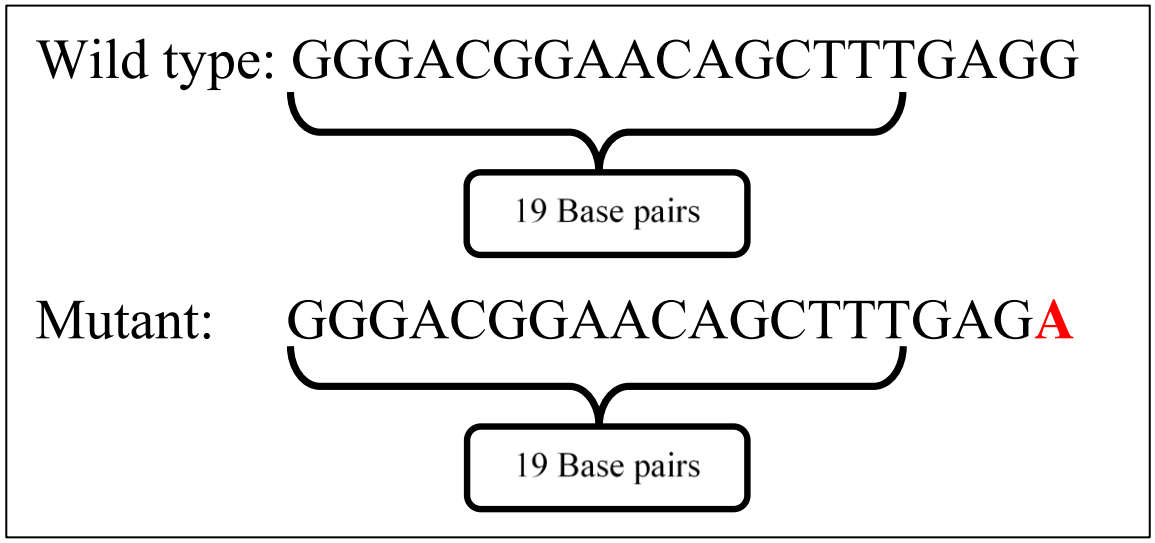


Figure 2-7. Example of forward primer for wild type and mutant of variant p.V272M *TP53*

The most appropriate common reverse primers are then calculated and displayed. The chosen primers were ordered from Eurofins Genomics at http://www.eurofinsgenomics.eu/ Figure 2-8.

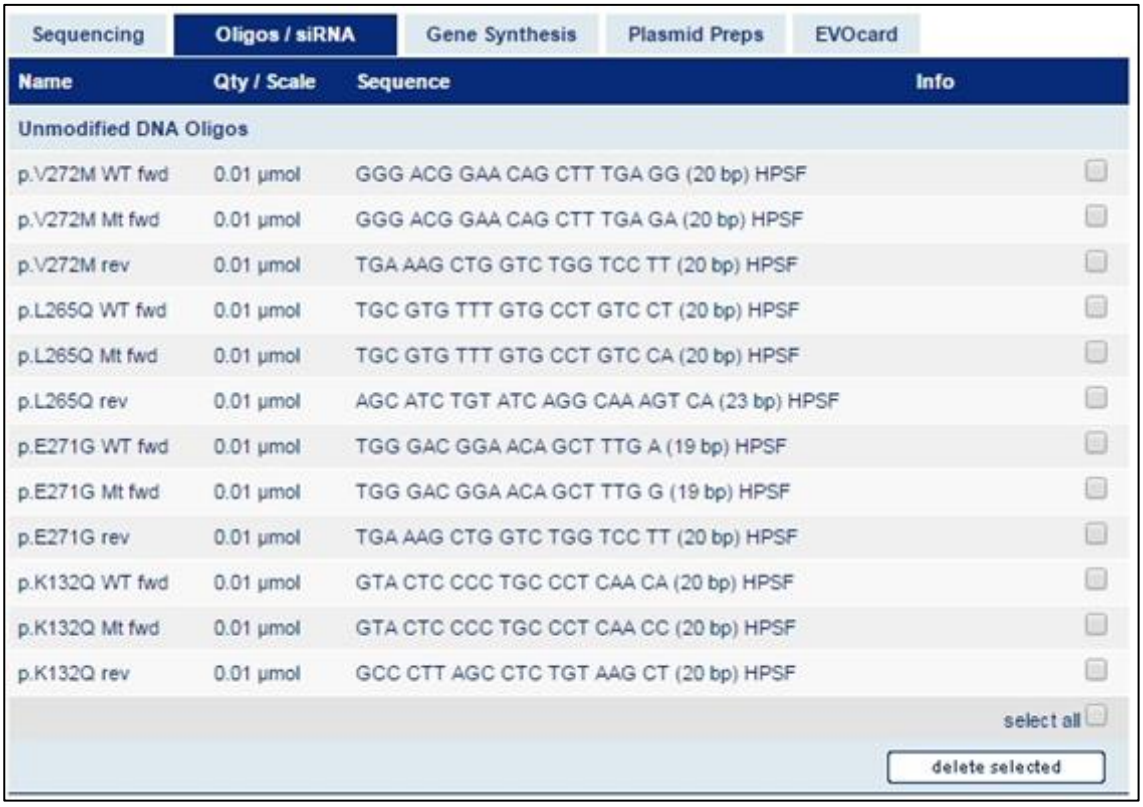


Figure 2-8. Screenshot of order page from Eurofins Genomics showing the wild type and mutant primers for 4 TP53 variants.

2.3.7 TP53 primer optimisation

Optimisation of the *TP53* primers was kindly performed by Miss Hollie Pufal, an undergraduate biochemistry student at York University under my supervision. In the absence of known cell lines harbouring the variants, positive controls were developed by performing quantitative PCR using the original DNA from which the variant was first identified with the respective mutant primers. The resulting PCR product was then run on a 1.8% agarose gel at 400mA for 40 minutes with an appropriate ladder. The region of gel corresponding to the expected product length was then cut out using a scalpel and DNA extracted using the methods previously described. This extracted DNA was then re-amplified by PCR with the same conditions. For negative controls DNA derived from PDAC cell lines used in chapter 4.2.1 was used.

The first stage of optimisation was to determine the most efficient annealing temperature for each primer by using temperature gradient PCR. The following annealing temperatures were investigated: 55.0°C, 58.2°C, 61.8°C and 63.7°C for primers p.V272M and p.L265Q and 68.8°C, 70.2°C, 73.4°C and 75.0°C for the p.E271G primer. For each 12.5µL reaction the following reagents were used: 1.25µL of

Gold Buffer (cat. no. P2317, Sigma, UK), 0.78µL of $MgCl_2$ (cat. no. M8787, Sigma, UK), 0.26 µL of dNTP mix (cat. no. dNTP100A, Sigma, UK), 0.26µL forward primer, 0.26µL reverse primer, 8.625µL of molecular grade DNA free water (cat. no. W4502-11, Sigma, UK) and 1µL of DNA. The PCR parameters were as follows: pre-incubation for 13 minutes at 95°C, DNA amplified for 45 cycles of 94°C for 50 seconds, the variable temperature for 15 seconds, 72°C for 1 minute, and finally 10 minutes at 72°C. The product was then run on a 1.8% agarose gel at 400mV for 40 minutes to assess product amplification. If a product of appropriate size was clearly amplified in the absence of negative control DNA bands, this annealing temperature was taken forward to quantitative PCR.

To optimise cycle number, minimise primer dimer formation and determine the optimal cut off temperature at which quantification is calculated various PCR parameters were tested. The Light cycler® 480 instrument (cat. no. 051015278001, Roche, UK) was used with 10µL reactions made up with 3.5µL molecular grade DNA free water, 5µL Roche Master Mix containing FastStart Taq DNA Polymerase, reaction buffer, dNTPs, MgCl2 and SYBR Green I dye (Cat. No. 04707516001, Roche, UK), 0.25µL of forward primer, 0.25µL of reverse primer and 1µL DNA. The PCR parameters were: 95°C for 13 seconds, 45 cycles of: 94°C for 50 seconds, various temperatures for 15 seconds, 72°C 1 minute and finally 72°C for 10 minutes. Various parameters were assessed by running PCR products on a 1.8% agarose gel at 400mV for 18 minutes with an appropriate ladder and then assessing amplification at the expected base pair length.

To test the primer specificity quantitative PCR was performed for various dilutions shown in Table 2-4. Each dilution was run on the Lightcycler® with the same reagents and the optimised PCR conditions. The recorded Cp values for both positive and negative controls were compared with that of the mutant.

Table 2-4. Dilutions of mutant and wild type DNA used to test specificity of primers

	Volume of Mt DNA (µL)	**Volume of Wt DNA (µL)**
Pure Wt	0	100
1 in 100	1	99
1 in 1,000	0.1	99.9
1 in 10,000	0.01	99.99
Pure Mt	100	0

2.3.8 TP53 NGS development

To assess the consistency and accuracy of the Ion torrent™ variant caller software a library preparation was duplicated on a separate chip. Variants and their reported frequencies were compared between the two runs.

2.3.9 FFPE tumour sampling

For each patient undergoing FFPE tumour analysis the full set of H&E stained slides including all sections of the resected specimen were microscopically reviewed by specialist histopathologist Dr Zainab Abdul-Rahman. Following examination of the slides, one slide was selected which was deemed to have the highest tumour cell content of the specimen. Upon this slide an area was marked with red pen corresponding again to the region of highest tumour cellularity as shown in Figure 2-10. For each sample the corresponding formalin fixed paraffin embedded (FFPE) block was obtained. Correlation was made between the marked area of interest on the selected slide and the FFPE block. The region of interest was then carefully excised with a scalpel macroscopically. This tumour tissue was then remounted into another paraffin embedded block. Five to ten 7 μm sections were cut from this block and loaded into an Eppendorf tube. DNA was eluted from these sections using the QIAamp DNA FFPE Tissue Kit (Cat. No. 56404, Qiagen, UK) according to the manufacture's guidelines. The eluted DNA was quantified using the Qubit 2.0 fluorimeter and diluted to 10 genomes/μL with molecular grade DNA free water (cat. no. W4502-11, Sigma, UK).

As mentioned in the introduction, NGS of primary tumour FFPE has some limitations and has not yet been shown to improve outcomes in PDAC. However, until the role of CTCs in PDAC has been clearly defined simultaneous sequencing of FFPE and CTCs may help guide future research projects. As such, in this thesis, some patients have both the primary tumour FFPE and enriched CTCs sequenced for comparision.

2.3.10 NGS set-up for cell culture

Blood from 5 patients (1,2,3,6 and 7) underwent cell culture by Dr Nick Bryan. The cells were trypsinised, washed and eluted using the MagNa pure system. The *TP53* gene was amplified by PCR by Dr Li Yan. The PCR parameters were as follows: pre-incubation for 13 minutes at 95°C, DNA amplified for 40 cycles of 94°C for 50

seconds, 54°C for 45 seconds, 72°C for 1 minute, and finally 10 minutes at 72°C. The reagents used were 2.5µL of AmpliTaq Gold® with Buffer (cat. no. P2317, Sigma, UK), 1.56µL of $MgCl_2$ (cat. no. M8787, Sigma, UK), 0.52 µL of dNTP mix (cat. no. dNTP100A, Sigma, UK), 0.4µL of forward primer, 0.4µL of reverse primer, 17.49µL of molecular grade DNA free water (cat. no. W4502-11, Sigma, UK) and 2µL of DNA. The single bands were excised from the gel and DNA was extracted using the Agarose Gel DNA Extraction Kit (cat. no. 11 696 505 001, Roche, UK) according to manufacturer's guidelines. The samples were sent to Source BioScience, Nottingham, UK for Sanger sequencing. KRAS analysis was performed by arms PCR also by Dr Li Yan as described in chapter one.

Cell culture of CTCs appears to be a very attractive option as it would overcome the main barrier to the utility of CTCs; the rarity of CTCs amongst vast quantities of WBC and would bypass the need for an enrichment process. Despite the attractiveness of this method successful culture of CTCs has not yet been achieved. Here, an attempt is made to culture CTCs and I describe the process to test whether it was successful.

2.4 Application of CTC technologies to PDAC cohorts

2.4.1 PDAC cell line mixing for Oncomine™ Solid Tumour DNA panel evaluation

The Oncomine™ panel contains multiplex primers for the hotspots of 22 frequency mutated genes. A list of the genes and their reported frequency in PDAC is listed in Table 2-5

Table 2-5 Full list of the 22 genes in the Oncomine™ panel and their reported frequency in PDAC

Mutation frequency of PDAC samples on Oncomine™ panel					
Gene	No. of samples tested	Samples with mutation	% in PDAC	Function	Chromosome
KRAS	6314	4265	67.55	Oncogene	12
TP53	2104	887	42.16	TS	17
SMAD4	1854	227	12.24	TS	18
CTNNB1	1654	131	7.92	Oncogene	3
PIK3CA	1499	29	1.93	Oncogene	3
ERBB4	1310	20	1.53	Oncogene	2
BRAF	1696	22	1.30	Oncogene	7
FBXW7	1361	15	1.10	TS	4
ERBB2	1373	15	1.09	Oncogene	17
NOTCH1	1352	11	0.81	Oncogene	9
STK11	1482	12	0.81	TS	19
PTEN	1389	11	0.79	TS	10
FGFR1	1318	7	0.53	Oncogene	8
NRAS	1453	7	0.48	Oncogene	1
EGFR	1781	7	0.39	Oncogene	7
MET	1380	5	0.36	Oncogene	7
DDR2	1287	4	0.31	Oncogene	1
MAP2K1	1314	3	0.23	Oncogene	15
FGFR3	1358	3	0.22	Oncogene	4
AKT1	1389	3	0.22	Oncogene	14
FGFR2	1620	3	0.19	Oncogene	10
ALK 2	0	0		Oncogene	2

Data was extracted from the COSMIC database. Tumour Suppressor (TS) (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/) 17th September 2015

Three PDAC derived cell lines: PANC-1 (cat. no. 87092802, Sigma-Aldrich, UK), MIA-Pa-Ca-2 (cat. no. 85062806, Sigma-Aldrich, UK) and SUIT-2 were cultured by Dr. Victoria Shaw in RPMI-1640 Medium (cat. No. R8758, Sigma-Aldrich, UK) supplemented with Foetal Bovine Serum (FBS, cat.no. 12003C, Sigma-Aldrich, UK) and glutamate. At approximately 80% confluency the cells were trypsinised, washed and pelleted by spinning at 1,000g for 10 minutes, the supernatant was removed and the cell pellet stored at -80^{o}C until used for DNA extraction. The pellet was then re-suspended in 400 µL PBS and DNA eluted using the MagNA Pure instrument previously described. The concentration of DNA was equalised by adding nuclease-free water (cat. No. AM9939, Ambion^{TM,} UK) as appropriate to all three to achieve a concentration of 100ng/µL using the Qubit 2.0 fluorimeter described previously. Different volumes of the cell line DNA were combined in the ratios shown in Figure 2-9 to make 11 separate samples. Mycoplasma testing was not routinely performed. Genomic sequencing of the cell line was performed at various dilutions using NGS and confirmed the expected mutations and purity of the cell line.

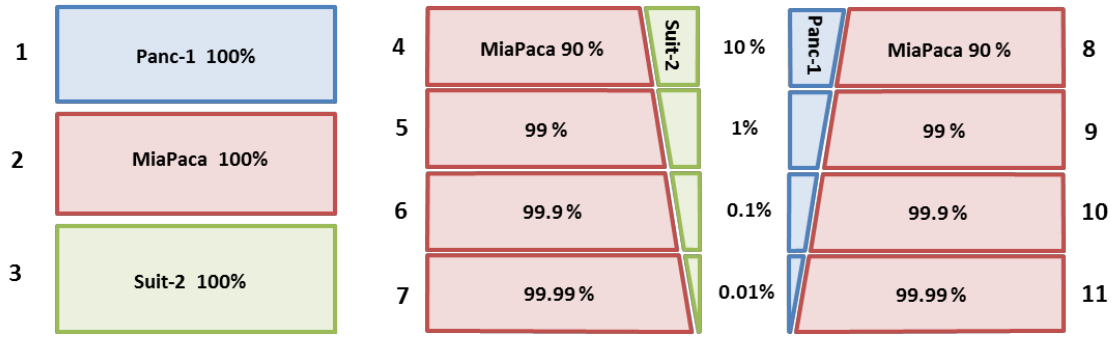


Figure 2-9. Various mixtures of cell lines used to make up 11 separate samples. Box sizes are not to scale

Each of the 11 samples, were then diluted to 5ng/µL and 2µL aliquots (10ng total DNA) of the cell line mix were used for library preparation using Oncomine™ Solid Tumour DNA kit (Life technologies, A26761, UK) in manner described later in this chapter.

2.4.2 Blood sample collection

All clinical blood samples were obtained by myself under the framework of the study: 'Comparison of techniques for detection of circulating tumour cells in peripheral blood', Research Ethics Committee reference: 08/H1011/36 with informed consent.

Two Sodium Heparin tubes 7.5 mL (cat. no. 01.1613.100, Sarstedt, UK) were used for blood sample collection in all patients. In the OncoQuik® cohort, blood was drawn on the morning of the day of surgery by the on-duty Clinical Research Fellow. The sample was then processed by technician Mr Neal Rimmer and stored within 4 hours as described in the section 4.2.3.1. Follow up samples were acquired when patients came for their postoperative check at the surgical outpatient clinic at the Royal Liverpool NHS Trust and processed in the same manner.
For the Rosette Sep™ cohort, patients were recruited at the Linda McCartney Oncology Unit in the Royal Liverpool University Hospital (RLUH) using the study framework. Blood was drawn immediately prior to administration of chemotherapy. The control samples were obtained from patients undergoing routine blood tests for non-cancer related conditions in the phlebotomy department of the RLUH Trust. A brief clinical history was taken to ensure there was no personal history of cancer. All samples in this cohort were processed directly with Rosette Sep™ within four hours, without freezing.

2.4.2.1 OncoQuik® blood sample processing
A two-step freezing protocol was used SOP GCLPTSS105/2 (Rimmer, Greenhalf, & Flaherty, 2014). In brief, after mixing by 10 blood tube inversions the tubes were centrifuged at 2,000g or 10 minutes. The plasma layer was removed and stored (green top cryovial) at -80°C. An equal amount of freeze media (mixture of 25% Dimethyl sulfoxide (DMSO) and 75% FBS) was added and the mixture re-suspended. The mixture was then aliquoted in up to 16 x 1.5mL blue top cryovial tubes and stored initially in a Mr Frosty container (cat. no. 5100-0001, ThermoFisher Scientific, UK) for 24-72 hours at -80°C and then transferred to -150°C for long-term storage.
For sample thawing the cryopreserved samples were incubated minimally for 10 minutes and maximally for 30 minutes in a water bath at 37°C. Washing buffer, 0.5% (w/v) bovine serum albumin (BSA, Sigma-Aldrich, UK) and phosphate buffered saline (PBS, Sigma-Aldrich, UK) mix was pre-warmed in water bath to 37°C. Once fully thawed all cryovial tubes (up to 16) were inverted twice to re-suspend cells and pooled into a 50 ml Falcon tube using a 1 mL pipette. Warmed washing buffer was added slowly at a rate of 1 mL per 5 seconds to equal volume, then more rapidly to make the final volume of 50mL. This was then centrifuged at 1398g for 10 minutes and the supernatant discarded (the purpose of this was to remove DMSO). The pellet was reconstituted with an equivalent volume of washing buffer to that from which it was

derived, e.g. 12 vials = 12mL of washing buffer. This was then layered in the upper compartment of the OncoQuik® tube (cat.no. 227250, Greiner bio-one, Germany) and proceeded as per the protocol described in chapter 3.2.4.

2.4.2.2 Rosette SepTM blood sample processing

The samples from this cohort were processed using 9mL of fresh whole blood divided into three aliquots of 3mL. 150 µL (50µL/mL) of the respective RS anti-body cocktail: RS-36, RS-45 and RS-56 were added to the 3mL blood aliquot and processed according to manufacturer's guidelines described in chapter 3.2.3. The pellet of enriched cells was reconstituted in 400 µL of PBS and used directly to elute DNA using the MagNA pure system described previously. DNA was quantified using the Qubit 2.0 fluorimeter and subsequently diluted to 20 genomes/µL.

2.4.3 Library preparation and sequencing

Library preparation was performed using the Oncomine™ Solid Tumour DNA kit (Life technologies, A26761, UK) according to manufacturer's guidelines. In brief there were 4 steps. Firstly, batches of 16 separate samples were used. DNA targets from each were amplified using PCR with the Oncomine™ Solid tumour DNA Panel primers and accompanying reagents provided. Secondly, the primer sequences were partially digested. Thirdly, adapters and barcodes were ligated onto the amplicons, using a separate barcode (1-16) for each sample Table 2-6. Finally, the library was equalised by first measuring the concentration of DNA in each barcoded sample with the Qubit 2.0 Fluorimeter. Each sample was then diluted as required with nuclease-free water. 10µL of each of the 16 equalised barcoded samples was then combined into a single Eppendorf tube and the concentration of the combined samples was again measured. NGS was again perfomed by technician Miss Katie Bullock on the Ion Personal Genome Machine® (PGM™) System (cat. no 4462921, Life Technologies, UK) according to manufacturer's guidelines and detailed in the SOP GCLPEQU022 - Use of the Ion Torrent™. The following kits and reagents were used: Ion PGM™ Hi-Q™ Sequencing Kit (cat. no. A25592, Life Technologies, UK), Ion PGM™ Template OT2 200 Kit (cat. no. 4480974, Life Technologies, UK), Ion Torrent 316 V2 chip (cat. no. 4483188, Life Technologies, UK)

Table 2-6. Sequence of barcode and adapter for each of the 16 barcodes used in the Oncomine™ Solid Tumour DNA kit

BARCODE	ADAPTER	SEQUENCE
1	GAT	CTAAGGTAAC
2	GAT	TTACAACCTC
3	GAT	CCTGCCATTCGC
4	GAT	TGGAGGACGGAC
5	GAT	TGAGCGGAAC
6	GAT	CCTTAGAGTTC
7	GAT	TCCTCGAATC
8	GAT	AACCTCATTC
9	GAT	CGGACAATGGC
10	GAT	TCCTGAATCTC
11	GAT	TAAGCCATTGTC
12	GAT	CTGAGTTCCGAC
13	GAT	CGGAAGAACCTC
14	GAT	TCTTACACAC
15	GAT	AAGGAATCGTC
16	GAT	TAGGTGGTTC

In previous chapters the concept of limiting dilution was described to distinguish genuine somatic variants from sequencing artefact. Dilutions of 10 genomes/µL were used in library preparation with between 5 to 10 sequencing replicates per each sample. Based on calculations of the typical mutation frequencies seen from OncoQuik® and Rosette Sep™ enriched blood in chapter three and on PDAC cell-line experiments in this chapter it was possible to lower the number of replicates and partially compensate for the reduction in total DNA by reducing the dilution. In both cohorts dilutions of 20 genomes/µL (40 genomes as 2µL per well) were used with two replicates/duplicates for each sample (80 genomes per sample in total). Such that an actual enriched sample

mutation frequency of 2.5% (2/80) would yield a measured 5% variant frequency from one of two duplicates. This arrangement was used where variants were required to be discovered a priori; in all OncoQuik® samples and for the first Rosette Sep™ (S1) samples. For the Rosette Sep™ follow up samples (S2 and S3) and all FFPE tumour samples tracking of the previously identified variants was required and so the standard 10 ng DNA was used without duplication according to manufacturer's guidelines.

2.4.4 FFPE Tumour sampling

The methods for FFPE tumour DNA extraction are described in chapter 2.3.9. In addition, where patients had a lymph node or an intraabdominal biopsy infiltrated with cancer this section of the slide was also marked Figure 2-10 and excised as described previously. The eluted DNA was quantified using the Qubit 2.0 fluorimeter and diluted to 5ng/µL with nuclease free water. 2µl or 10ng of DNA was used in the library preparation using only one barcode per sample.

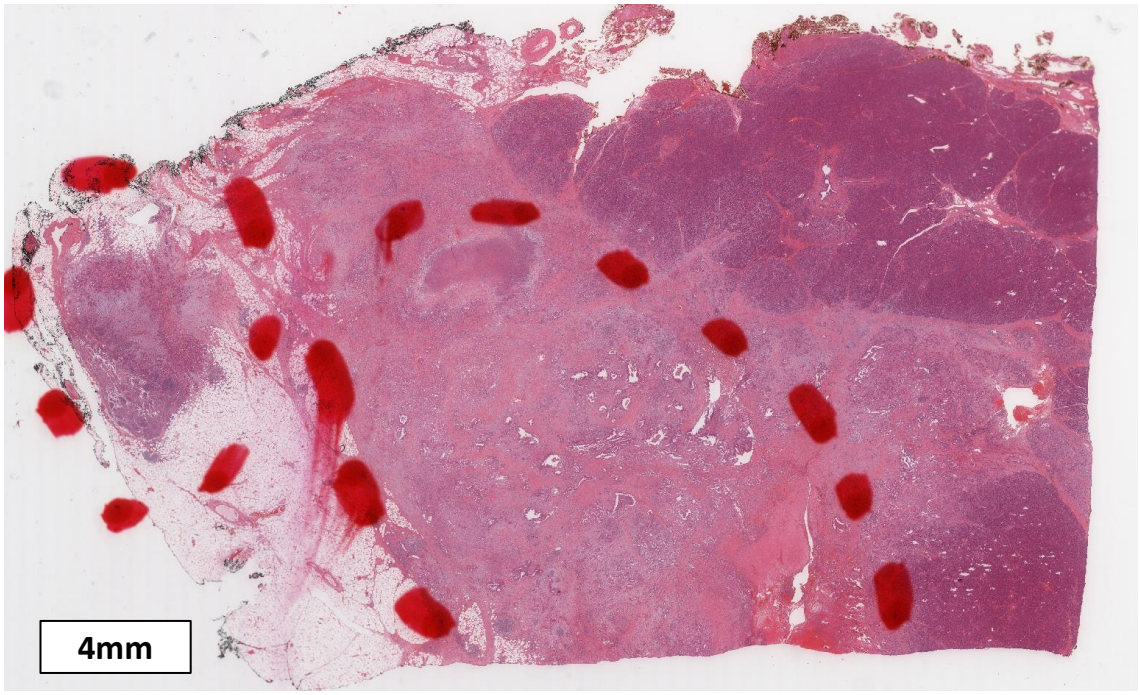


Figure 2-10. Tumour identification in primary and lymph node. The largest red dotted area represents highest tumour cellularity of primary PDAC in patient 5. The smaller red dotted area identifies involved lymph node in the same slide. Scale bar = 4mm

2.4.5 Variant characterisation

The Ion Reporter™ Software 5.0 bioinformatics tool available at https://ionreporter.thermofisher.com/ir was used. The human genome version 19 was used as the reference against which variants were called. Variant assessment was performed in 3 stages:

1. Ion reporter™ 5.0 filters:
    a. Minor Allele Frequency (MAF): <0.3
    b. Variant Effect: exclude 'refAllele' and 'synonymous'
    c. Variant Type: exclude 'REF' and 'NOCALL'
    d. P value: < 0.1

The P value represents the probability that the variant call is incorrect and is a logarithmic transformation of the Phred quality score value discussed below. The conventional value of P < 0.1 was used here to reduce the number of variants included due to errors in ion torrent base calling and alignment.

2. Post analysis assessment of scores:
    a. Sorting Intolerant from Tolerant (SIFT)
    b. Phred quality score

Filtered variants were then scored using the SIFT score which uses sequence homology to predict whether non-synonymous variants affect protein function and the Phred quality score Q. Q is defined as a property which is logarithmically related to the base-calling error probabilities P (Ewing 1998) where:

$$Q = -10\ log_{10}P$$

or

$$P = 10^{\left(\frac{-Q}{10}\right)}$$

such that a Phred quality score of 60 gives the chance of that base being incorrectly called as 1 in 1,000,000.

Finally, variants were cross-referenced with three variant databases.

3. Cross referencing with databases:
    a. Catalogue of Somatic Mutations (COSMIC)
    b. International Agency for Research on Cancer (IARC) TP53 database
    c. Single Nucleotide Polymorphism Database (dbSNP)

Presence in the Catalogue of Somatic Mutations (COSMIC) database confirms that the variant has previously been implicated as a somatic mutation in cancer (mutation present in tumour DNA and absent from matched normal DNA). For TP53 mutations the International Agency for Research on Cancer (IARC) TP53 database available at http://p53.iarc.fr/ compiles all reported TP53 mutations in sporadic cancers and describes, where available the expected tumour phenotype and functional and structural impact of mutations. Finally, the Single Nucleotide Polymorphism Database (dbSNP) compiles all reported germline polymorphisms and their minor allele frequency available at http://www.ncbi.nlm.nih.gov/SNP/.

Where a visual representation of the variant was required the intergrative genomic viewer (IGV) was used. Variants were excluded where there was clear evidence of mall-alignment or base calling errors particularly in insertion and deletion (indel) variants identified in higher polymer regions.

2.4.6 Statistical analysis

Differences in mutation frequency were tested between groups using, Pearson's chi squared and Fisher's exact tests as appropriate. Kaplan-Meier curves with log-rank test were used to compare outcome data such as survival.

2.4.7 List of Reagents

The list of reagents used, how they are made and their supplier are shown in Table 2-7.

Table 2-7 List of reagents used and their supplier

Reagent	Dilution/preparation	Supplier	Catalogue No.
Secretin	1 IU/Kg	Sanochemia, Germany	111527
AmpliTaq Gold® with Buffer	Undiluted	Sigma-Aldrich, UK	P2317
DNA Molecular Weight Marker VIII	Undiluted	Roche, UK	
dNTP mix	Undiluted	Sigma-Aldrich, UK	dNTP100A
Foetal Bovine Serum (FBS)	dissolved in 1L of deionized water	Sigma-Aldrich, UK	12003C
Ion PGM™ Hi-Q™ Sequencing Kit	Undiluted	Life technologies, UK	A25592
Ion PGM™ Template OT2 200 Kit	Undiluted	Life technologies, UK	4480974
Ion Torrent 316 V2 chip	Undiluted	Life Technologies, UK	4483188
Lymphoprep	Undiluted	Stem cell technologies	7811
$MgCl_2$	Undiluted	Sigma-Aldrich, UK	M8787
MIA-Pa-Ca-2 cell line	Undiluted	Sigma-Aldrich, UK	85062806
Molecular grade DNA free water	Undiluted	Sigma-Aldrich, UK	W4502-11
Oncomine™ Solid Tumour DNA kit	Undiluted	Life technologies, UK	A26761
PANC-1 cell line	Undiluted	Sigma-Aldrich, UK	87092802
Phosphate Buffered Saline (PBS)	dissolved in 1L of deionized water	Sigma-Aldrich, UK	P3818
RosetteSep™ CTC Enrichment Cocktail Containing Anti-CD 36	Undiluted	Stemcell technologies, UK	15167
RosetteSep™ CTC Enrichment Cocktail Containing Anti-CD 45	Undiluted	Stemcell technologies, UK	15122
RosetteSep™ CTC Enrichment Cocktail Containing Anti-CD 56	Undiluted	Stemcell technologies, UK	15177
RPMI-1640 Medium	Undiluted	Sigma-Aldrich, UK	R8758
SYBR Green I dye	Undiluted	Roche, UK	4707516001
Thermo Scientific™ 6X Orange DNA Loading Dye	Dilute 1:6 with water	Thermo Fisher, UK	R0631

3 PANCREATIC JUICE AS A POTENTIAL SOURCE OF BIOMARKERS FOR PDAC

3.1 Introduction

The collection of pancreatic juice by endoscopic retrograde cholangiopancreatography (ERCP) and direct cannulation of the main pancreatic duct risks post-ERCP acute pancreatitis (PEP)(Freeman ML, DiSario JA, Nelson DB, 2001). An alternative method of pancreatic juice collection from the duodenum and three different components of pancreatic juice are trialled, to determine if their sensitivity and specificity for PDAC is maintained

3.2 Results

3.2.1. Comparison of pancreatic juice supernatant vs pellet for biospecimen use

Pancreatic juice, analysed in both pellet and supernatant forms was collected from 35 patients with PDAC, 23 with chronic pancreatitis and 108 patients with gallstones (controls). The patient demographics are shown in Table 3-1 and the outcome of the molecular analysis according to group is shown in Table 3-2.

Table 3-1. Patient demographics of patients in SSECRETIN study

		Cancer n=35	**CP n=23**	**Control n=108**	**p value**	
Age	Median (IQR)	71 (63-75.5)	60 (42-67)	66 (56-77.3)	0.0014	(Cancer vs CP)
					0.0169	(Control vs CP)
					0.171	(Cancer vs Control)
Bilirubin	Median (IQR)	219 (64-316)	9 (5.5-13)	13.5 (6-44.5)	<0.001	(Cancer vs CP)
					0.0295	(Control vs CP)
					<0.001	(Cancer vs Control)

Table 3-2. Molecular analysis of the supernatant and pellet of pancreatic juice

	Supernatant				**Pellet**			
	n	% *CDKN2a* meth, median (IQR)	n	*KRAS* mutant: wild type (% Mutant)	n	% *CDKN2a* meth, median (IQR)	n	*KRAS* mutant: wild type (% Mutant)
PDAC	34	7.7 (0.5-30.2)	35	14:21 (40.0)	24	22.9 (4.31-72.3)	26	16:10 (61.5)
CP	21	6.01 (0.88-52.3)	23	10:13 (43.5)	20	9.75 (1.83-21.4)	21	8:13 (38.1)
Control	107	27.4 (15.3-38.5)	106	64:42 (60.4)	38	4.72 (1.02-16.7)	35	9:36 (20.0)

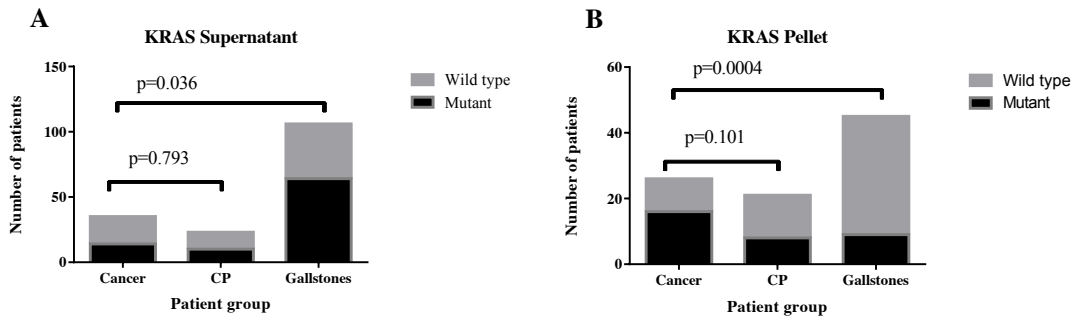


Figure 3-1. Mutant *KRAS* was significantly more frequent in PDAC patients compared to controls in pellet but not supernatant. P value calculated by Chi-square.

The supernatant samples performed badly in discriminating cancer from controls. In fact, the supernatant control samples had significantly more *KRAS* mutant (60% vs 40%, p=0.036) Figure 3-1A and higher % *CDKN2a* promoter methylation (27.4 vs 7.7, p=0.031) than supernatant cancer samples Figure 3-3 and Table 3-4. The pancreas juice pellet proved a better biospecimen with both KRAS and % *CDKN2a* promoter methylation significantly discriminating cancer from controls (p=0.004 and p=0.015 respectively) and trending towards significance in discriminating from cancer from CP (p=0.101 and p=0.051 respectively).

In previous work the optimal cut-off for % *CDKN2a* promoter methylation was determined to be 12%. This threshold was applied to the cohort, the resulting sensitivities and specificities in addition to *KRAS* is shown in Figure 3-2. In Figure Figure 3-2 three separate specificity values are shown. The ability to discriminate cancer samples from just the control samples, excluding CP is shown as Specificity (Control). Specificity (CP) represents the discrimination of cancer from CP alone (excluding control samples). And finally, Specificity (Control and CP) represents the ability to discriminate cancer samples from both control and CP samples combined.

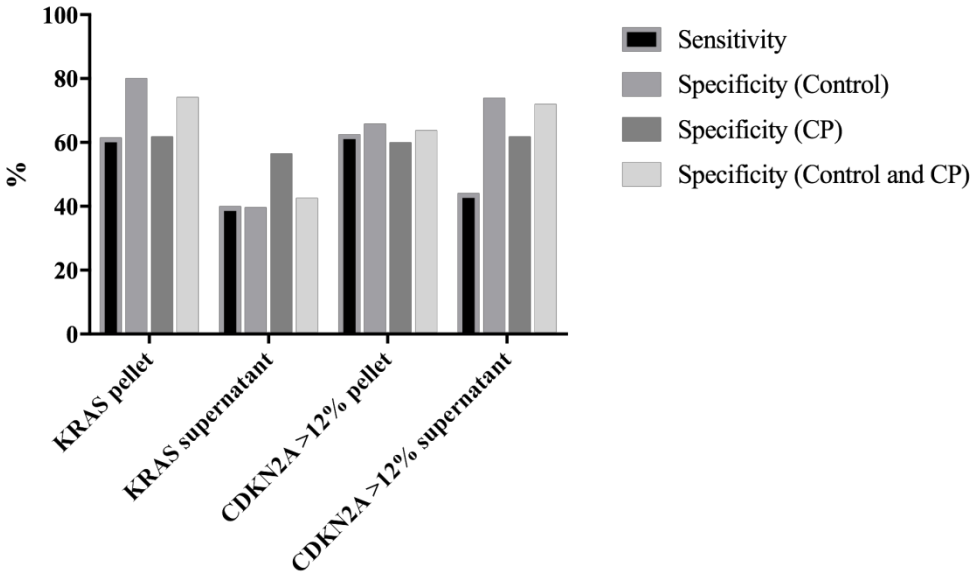


Figure 3-2. Sensitivity and specificity of KRAS and % *CDKN2a* methylation in supernatant and pellet.

The % of *CDKN2A* promoter methylation was significantly elevated in caner patients compared to control patients (P=0.015) and trended towards significance in discriminating from CP patients. It was surprising to observe that % of *CDKN2A* promoter methylation was significantly elevated in control patients compared to cancer patients (p=0.031)Figure 3-3.

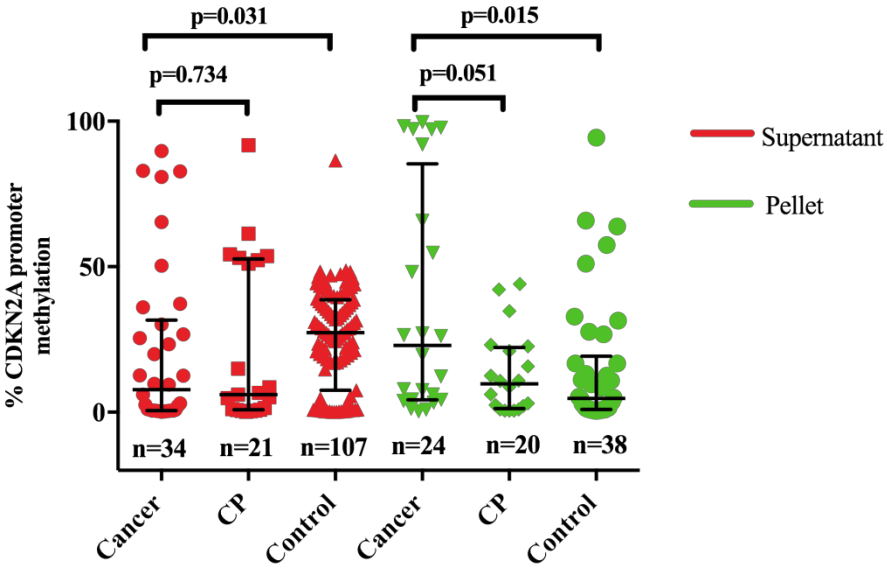


Figure 3-3. % *CDKN2a* promoter methylation is significantly elevated in cancer compared to control groups in pellet, but the reverse association is seen in supernatant. The ability to discriminate CP from cancer in the pellet trended towards significance. Error bars display median and inter-quartile range, p values calculated by Mann Whitney U.

The clinical utility of % *CDKN2a* promoter methylation alone in the pellet of pancreatic juice was limited with a relatively low area under the curve (AUC) of 0.684 in discriminating PDAC from controls blue line Figure 3-4. In an attempt to improve the discriminatory value of the test, the analysis was repeated with samples which were only *KRAS* mutant (green line Figure 3-4) and only KRAS The AUC was improved when combined with samples which were *KRAS* mutant (AUC=0.796)

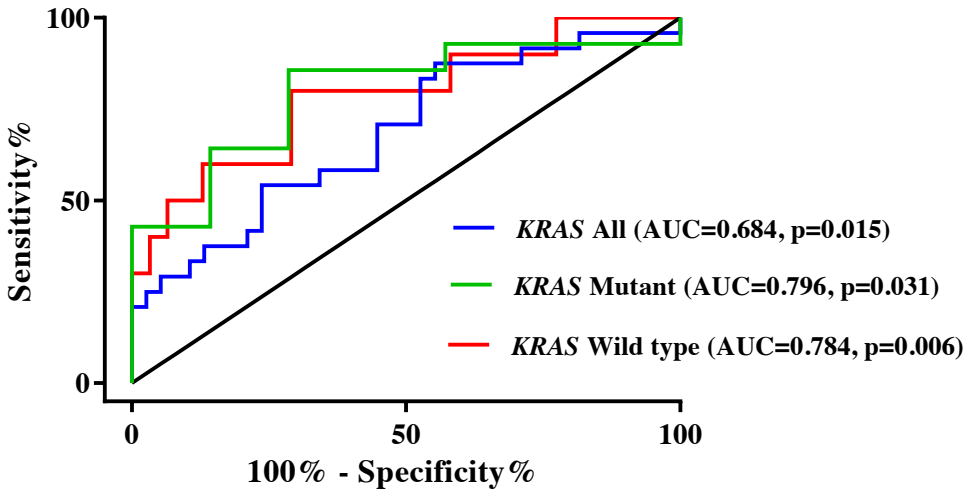


Figure 3-4. Receiver operating characteristic (ROC) curve analysis of % *CDKN2a* promoter methylation in discriminating PDAC from controls (no CP). Blue line shows all samples, PDAC n=24, controls n=38. Green line shows *KRAS* mutant only, PDAC n=14, controls n=7. Red line shows wild type *KRAS* only, PDAC n==10 and controls n=31

3.2.2. Comparison of the molecular profile of whole pancreatic juice to the pellet and supernatant

In previous the sections we have seen the results obtained using the supernatant and pellet of pancreatic juice. Since this time, the Goggins group have shown improved results using whole juice. (Kanda, et al., 2013) We didn't have whole juice samples for the previous dataset as the samples were spun down directly into supernatant and pellet immediately after collection. We therefore conducted a new study prospectively collecting secretin stimulated pancreatic juice from the duodenum as before, but with the aspirate split into two; the first part spun into supernatant and pellet, and the second part stored as whole juice.

To date in this on-going study 7 patients have been recruited to the PDAC group, and 28 to the control group (consisting of 25 patients with gallstones and 3 with chronic pancreatitis). For whole juice the median % CDKN2a promoter methylation was higher in the cancer group, this did not reach significance, but in this relatively small group of patients no significance was seen with pellet either (which we saw in the analysis above was significant with larger numbers). In this small group supernatant had a slightly higher median level of methylation with cancer than with controls, but again this was

not significant Figure 3-5. Therefore, whole pancreatic juice may be better than pellet or supernatant at differentiating cancer, but the difference is not very great, for individual samples there is not a great benefit of using whole juice rather than pellet Figure 3-6.

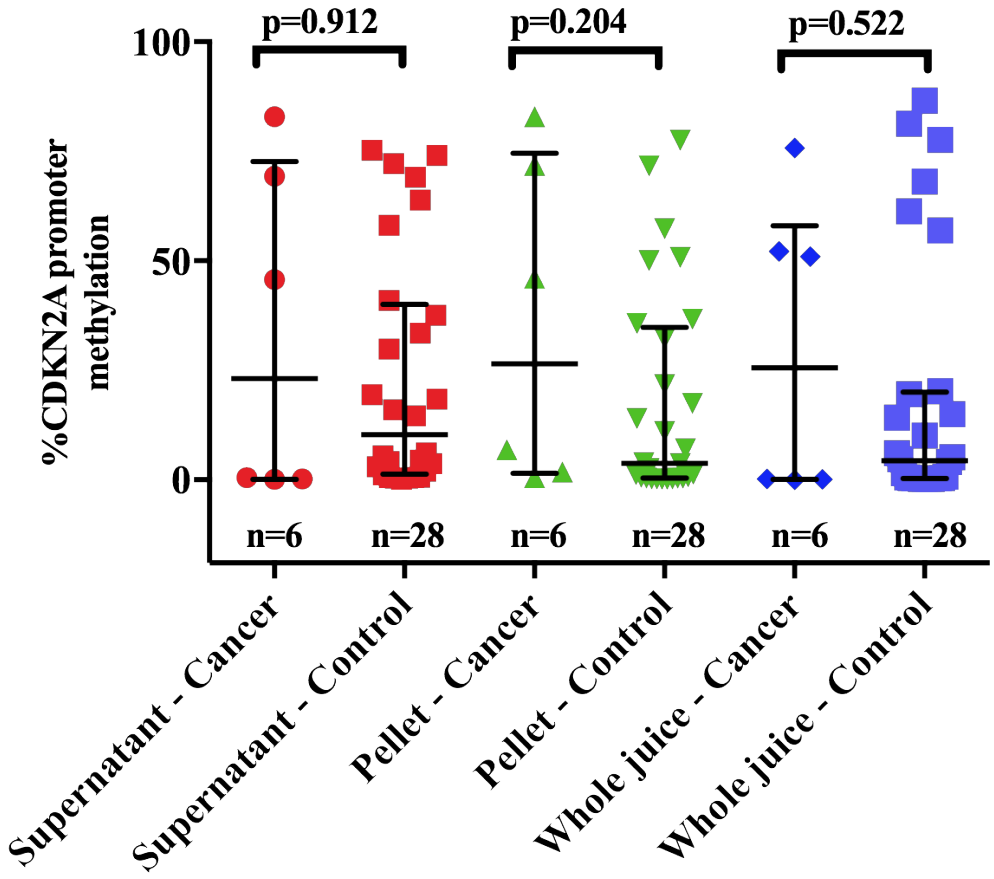


Figure 3-5. *% CDKN2a* promoter methylation is not significantly elevated in cancer compared to control groups in supernatant, pellet or whole pancreatic juice. Error bars display median and inter-quartile range, p values calculated by Mann Whitney U. *One cancer sample not yet had %CDKN2A promotor methylation analysed and so excluded

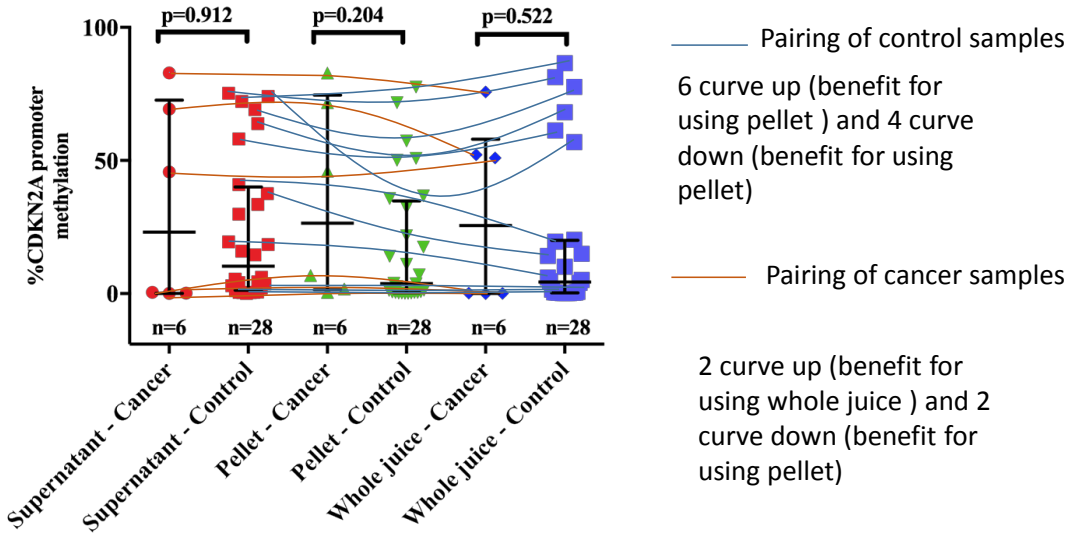


Figure 3-6. Matched samples from Figure 3-5 are linked showing that there is no significant trend for cancer pellet to have lower methylation than in whole juice from cancer (or vice versa for controls)

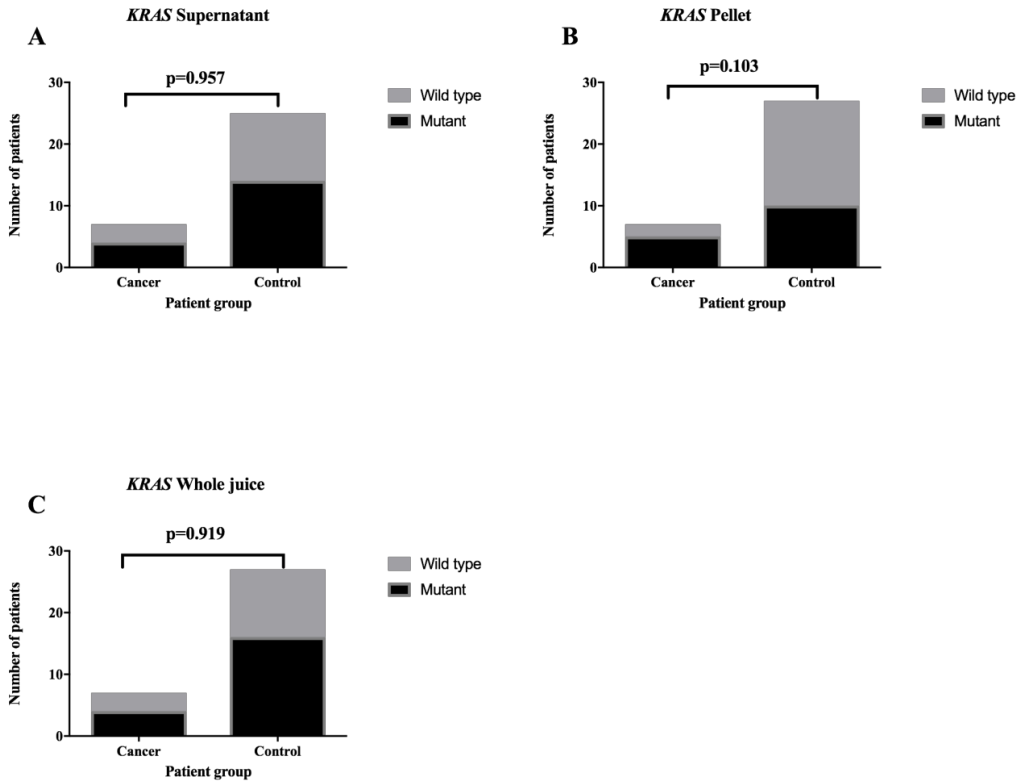


Figure 3-7. Mutant *KRAS* was not significantly more frequent in whole juice from cancer compared to controls.

As in the previous section sensitivity and specificity was again low for *KRAS* mutation status in supernatant (57.1% and 44.0% respectively), pellet (71.4% and 63.0% respectively) and whole juice (57.1%, 40.7% respectively)Figure 3-7.

Though the *KRAS* mutation in the pellet trended towards significant association with cancer, p=0.103, none of the sample types reached significance Table 3-3.

Table 3-3. KRAS mutation was not significantly associated with cancer in any of the sample types. P value calculated by Chi-square.

Pancreatic juice component	**KRAS**	**Cancer**	**Control**	**Total**	**p value**
Supernatant	Mutant	4	14	18	0.819
	Wild type	3	13	14	
	Total	7	27*	32	
Pellet	Mutant	5	10	15	0.103
	Wild type	2	17	19	
	Total	7	27*	34	
Whole juice	Mutant	4	16	20	0.919
	Wild type	3	11	14	
	Total	7	27*	34	

*One control patient excluded as did not have *KRAS* mutation performed.

3.3 Discussion

The retroperitoneal position of the pancreas and the low resection rates of PDAC make acquisition of a biospecimen for PDAC challenging(Speelman et al., 2015). Pancreatic juice is an attractive option as PDAC arises from ductal cells, shed tumour cells are therefore likely to be contained within the juice which can be obtained endoscopically. Indeed, it has been shown that molecular analysis of pancreatic juice can be used to stratify risk of PDAC(Yan et al., 2005). ERCP however is invasive and is associated with a significant risk of PEP(Cheng et al., 2006). Studies from our group have shown that prophylaxis (stent and diclofenac) can reduce the risk of PEP(Nicholson et al.,

2015). Despite this, the risk is still too high for routine clinical use. A modification to the collection technique (collection from the duodenum) essentially eliminates the risk of PEP associated with pancreatic duct cannulation but compromises the biospecimen. Of the three components of pancreatic juice, the cell pellet was superior to supernatant in discriminating PDAC from control samples and no evidence was found that pellet was inferior to whole juice. KRAS mutation status was superior to the percentage of *CDKN2a* methylation. The sensitivity and specificity of *KRAS* status for PDAC in the pellet of pancreatic juice was 83.3% and 64.3% respectively, but was still insufficient for routine use.

It is known that iron-containing proteins and their breakdown products, such as bilirubin, and bile salts are major inhibitors of PCR which may affect assays in some jaundice patients when performed on blood, stool and urine where by-products of bilirubin are found(Schrader, Schielke, Ellerbroek, & Johne, 2012). However, there has been no report to date of the serum bilirubin interference on PCR assays performed on pancreatic juice despite many groups using this biospecimen(Yan et al., 2005)(Mitsuro Kanda, Sadakari, et al., 2013)(Fukushima et al., 2003). No modification to the PCR assays on pancreatic juice were therefore used.

Despite the described technique modification avoiding pancreatic duct cannulation, endoscopy in itself is still an invasive and expensive investigation requiring specialist equipment and skilled endoscopists. Serial sampling in this instance is not feasible. The use of blood as a biospecimen in this respect has many obvious advantages over pancreatic juice in that it is relatively non-invasive, cheap, safe and permits serial sampling. It is also a route by which PDAC along with most other cancers metastasise and therefore potentially offers a more complete picture of the systemic disease than just the primary tumour alone. The remaining chapters evaluate blood as a biospecimen.

4 NEGATIVE DEPLETION APPROACH TO USING BLOOD AS A BIOSPECIMEN

4.1 Introduction

Whole blood allows minimally invasive serial sampling and is therefore an attractive biospceimen in PDAC given the biological and physical flaws associated with interrogation of solid primary tumour. Here, both an in-house method of CTC enrichment using negative depletion and the sensitivity of low frequency variant detection using next generation sequencing (NGS) and limiting dilution is evaluated.

4.2 Results

4.2.1 A single CD45 depletion is optimal in small volume spiked samples

The enrichment capabilities of a CD45 depletion was first investigated using PANC-1 cells spiked into peripheral blood mononuclear cells (PBMCs) thus removing the variable of other blood constituents such as platelets, red blood cells and polymorphonuclear white blood cells. Following one CD45 depletion the mean Cp value (n=3) of wild type DNA increased from 26.8 (SD=0.71) to 32.7 (SD=0.89) p=0.02 Figure 4-1A representing a 59.3 (2 x $10^{5.89}$) fold reduction in wild type DNA. There was an increase in mutant (p.G12D, PANC-1) DNA concentration observed with the mean (n=3) Cp value falling from 29.4 (SD=0.38) to 27.6 (SD=0.56) p=0.02 Figure 4-1A equating to a 3.4 (2 x $10^{1.8}$) fold enrichment. This suggests that the CD45 antibody and bead complex are specifically binding PBMCs but not PANC-1 cells, as expected. Further CD45 depletions did not lead to further reductions of wild type DNA p=0.592 or enrichment of mutant DNA p=0.135 Figure 4-1A. Indeed, three CD45 depletions removed the mutant signal altogether, suggesting serial depletions eventually have a non-specific binding effect depleting PANC-1 cells as well.

Next, the experiment was repeated using whole blood incorporating the additional factors of red blood cells and polymorphonuclear white blood cells. The whole blood was diluted 1:10 with PBS to remove the variable of blood viscosity, in particular the ability of the large ferrous beads to traverse through densely packed RBC. Again, a 74.0-fold (2 x $10^{6.2}$) reduction in wild type DNA was achieved with 1 x CD45 and 1 x RBC depletions from a Cp of 26.4 (n=3) to 32.61 (n=3) p=0.007 Figure 4-1B with a corresponding 3.5-fold (2 x$10^{1.8}$) enrichment of mutant DNA from mean Cp value 30.0 (n=3) to 28.2 (n=3) p=0.011 Figure 4-1B. Additional RBC depletions further reduced

the wild type DNA on both the second (P=0.0629) and third depletions (P=0.0265) but the mutant signal was not enriched on the second (P=0.920) and completely removed on the third Figure 4-1B. Thus, a single CD45 depletion appeared optimal in both PMBCs and in diluted whole blood, whilst RBC depletion appeared to have non-specific binding. We next investigated, whether single or multiple CD45 depletions were optimal in whole blood and whether the addition of a RBC depletion to a single CD45 depletion improved enrichment in whole blood. Neither the addition of a second, nor a third CD45 depletion improved the enrichment of the mutant signal of various numbers of PANC-1 cells spiked into 30μL of undiluted whole blood.

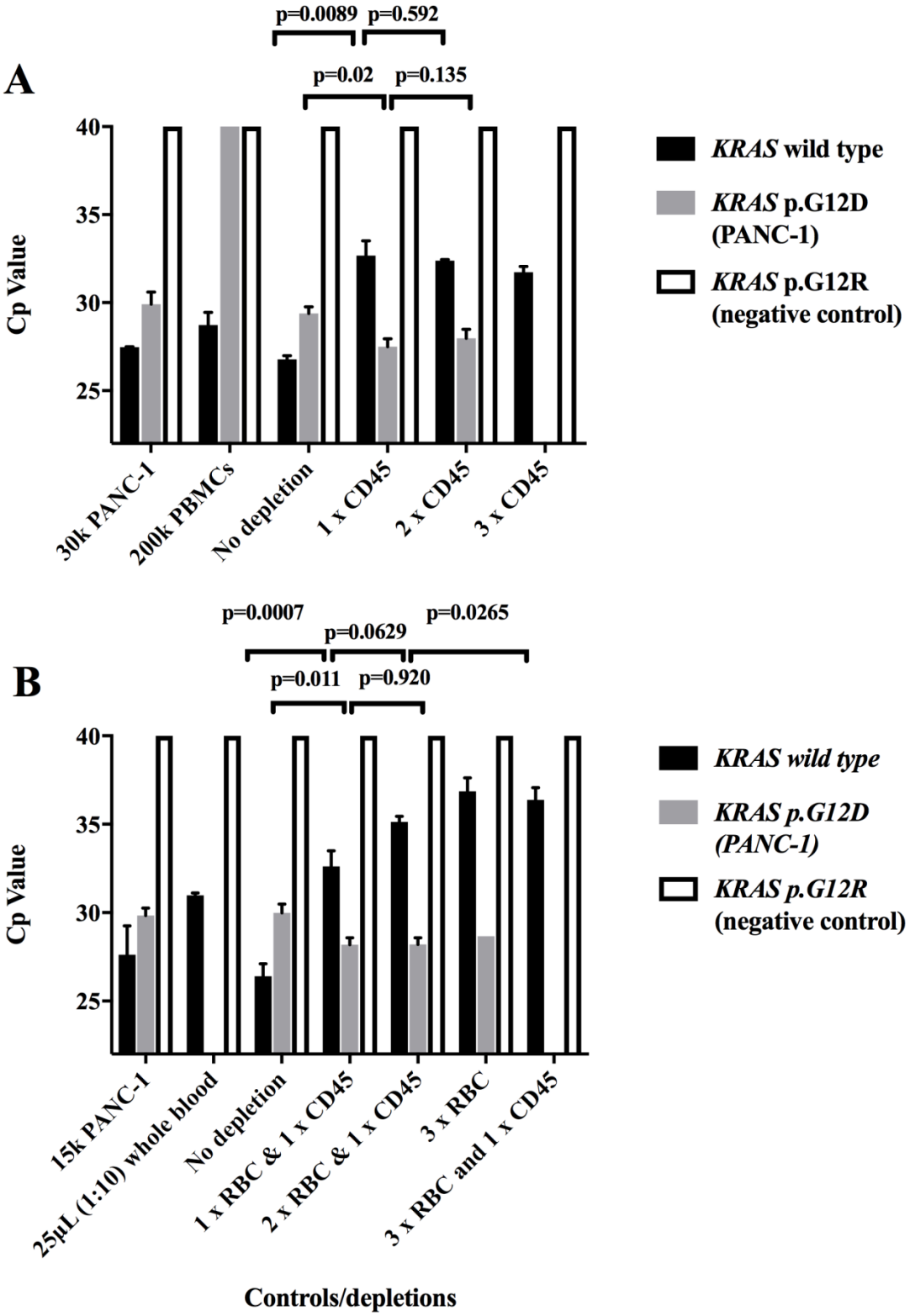


Figure 4-1 CD45 depletions significantly reduce wild type DNA and enrich PANC-1 signal in both PBMCs and (1:10) diluted whole blood experiments. Subsequent depletions reduce and eliminate PANC-1 signal altogether A) 30k PANC-1 cells spiked into 200k PBMCs. B) 15k

The addition of a RBC depletion to a single CD45 depletion also did not have a significant effect on the enrichment of the mutant signal in any of the experimental set-ups Figure 4-2A-D.

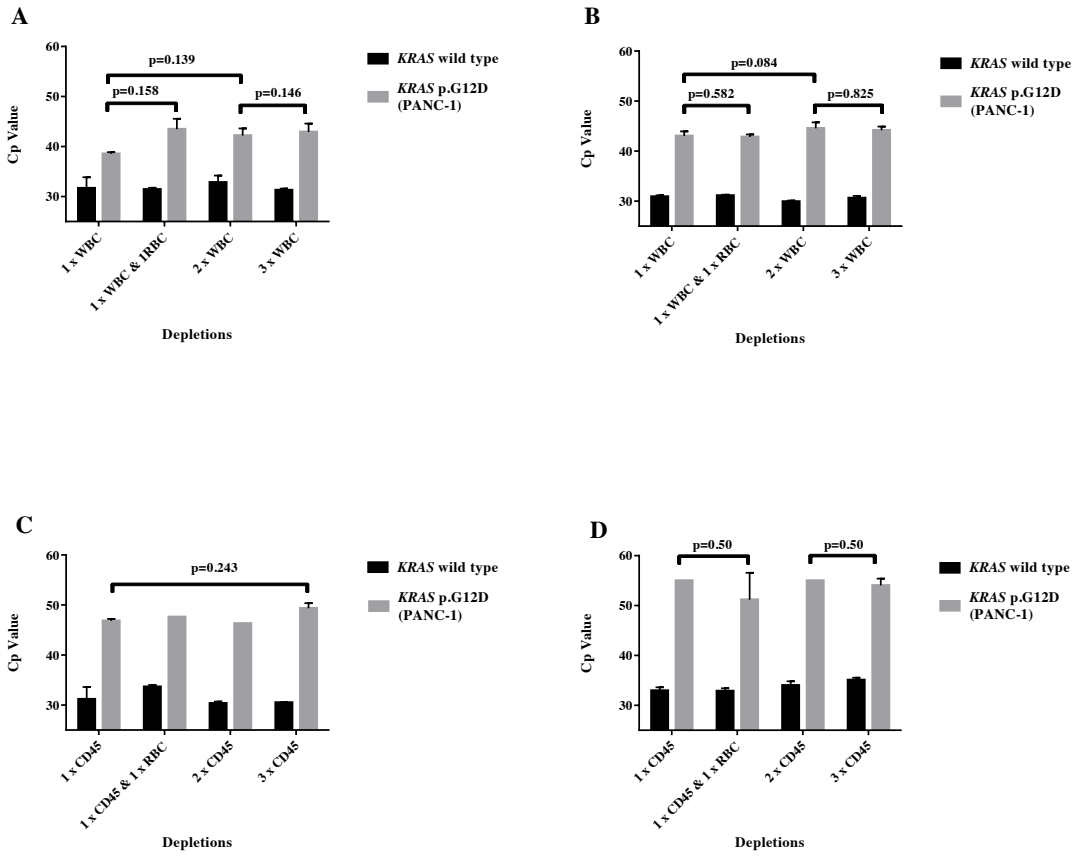


Figure 4-2. Increasing the number of CD45 depletions beyond the first does not improve PANC-1 enrichment when spiked into 30μL of whole blood. The addition of a RBC depletion to the first CD45 depletion also did not improve enrichment. Various quantities of PANC-1 cells A) 10,000, B) 1,000, C) 110, D) 8, were spiked into 30μL of whole blood. P values were derived from paired t test, error bars display mean and standard deviation.

4.2.2. Increases in both antibody and bead volume improve efficiency of CD45 depletions

Having defined the optimal number and type of depletions (1 x CD45) the efficiency of this depletion in terms of bead and antibody volume was investigated. Liner regression analysis showed a strong relationship between antibody volume and amount of DNA removed from the sample attached to the beads. This was true for both 80µL (R^2=0.949, p<0.0001) and 40µL (R^2=0.703, p=0.010) bead volume variable Figure 4-3A. In addition, using 80µL of beads removed significantly more DNA from the samples than using 40µL of beads (p=0.0251) Figure 4-3A. As expected, the amount of DNA remaining in the supernatant of the sample was greater when using the smaller volume (40µL) of beads though this did not reach statistical significance (p=0.223) Figure 4-3B. An unexpected relationship was observed between increasing volume of antibody and increasing amount of DNA remaining in the supernatant with both 40µL of beads (R^2=0.530, p=0.0021) and 80µL of beads (R^2=0.581, p=0.0025), this was a significant relationship albeit weak. This relationship however was not as strong as that seen with DNA removed with the beads and may, in part, be due to the greater proportion of DNA remaining behind in the sample. An explanation for this may be that volume was lost in the supernatant and therefore the sample contained more cells per microliter.

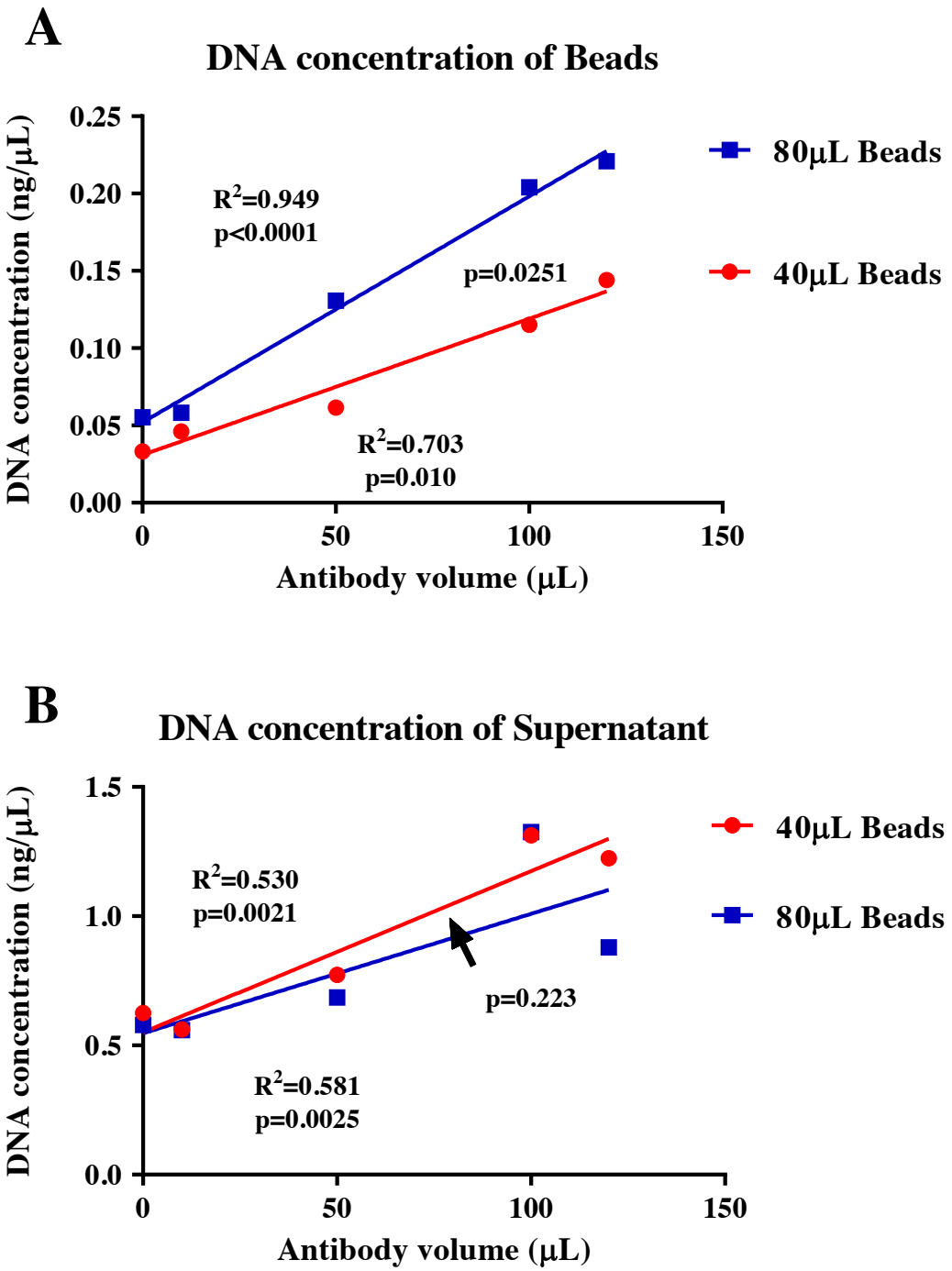


Figure 4-3. Liner regression analysis shows a strong relationship between antibody volume and amount of DNA removed from the sample for both 80μL and 40μL volume of beads. A) Concentration of DNA attached to the beads removed from the sample. B) Concentration of DNA remaining in the supernatant after removal of beads.

4.2.3. Ion torrent™ is sensitive down to a variant frequency of 0.1% where coverage is greater than 500

Before proceeding with spiked samples into clinical blood volumes (>7mL) the sensitivity of the Ion Torrent™ NGS platform in conjunction with limiting dilution was assessed Table 4-1 and demonstrates that the Ion Torrent™ base calling software accurately identified the serial 10-fold dilutions in the PANC-1 cell line. It can be seen that where the total coverage is below 500, the ability to identify very low frequency variants is diminished . It can be seen that the total percentage of mutant sequence in the Pure PANC-1 sample was only around 70%. Clearly some PANC-1 cells contain some wild type TP53 sequence. This finding has not previously been reported in the literature. It is known that PANC-1 is polyploid and our data indicates that 1 chromosome 17 out of 5 may have a wild type sequence. Conventional Sanger sequencing is not accurate below 25% which may explain why this has not been reported previously. Note that the p.R273H mutation is dominant, hence 1 wild type copy in 5 would still give a mutant phenotype.

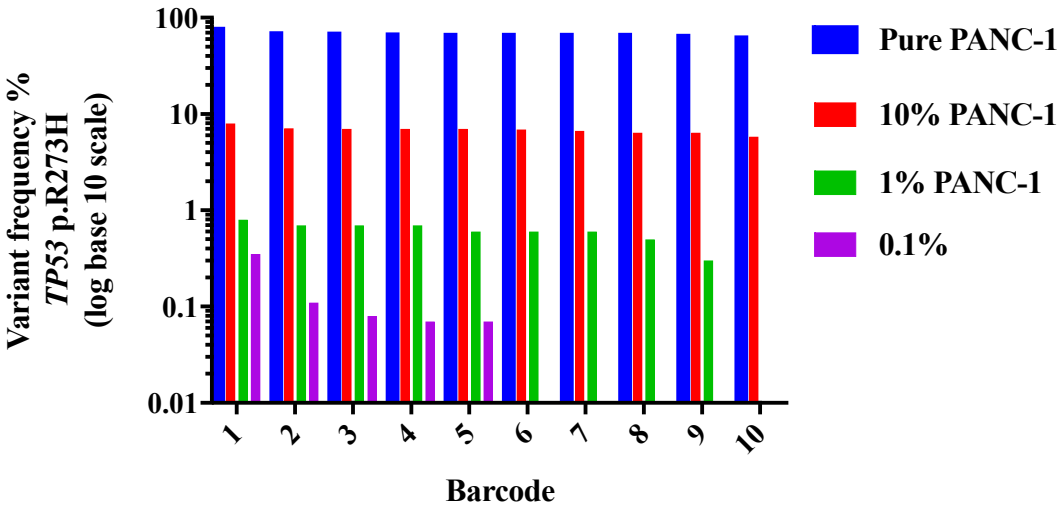


Figure 4-4 The Ion Torrent™ PGM platform identifies the appropriate variant frequency for PANC-1 cell lines serially diluted 10-fold down to 0.1%. Y axis is log base 10 scale

Table 4-1 Frequency of TP53 variant p.R173H (PANC-1) reported by Ion Torrent™ for serial 10-fold dilutions of PANC-1 DNA. Each barcode represents a sampling replicate, each diluted to 10 genomes/μL

	Pure PANC-1			**10% PANC-1**			**1% PANC-1**			**0.1% PANC-1**		
barcode	Variant	Total	%	Variant	Total	%	Variant	Total	%	Variant	Total	%
1	89	110	80.9	241	3013	8	4	493	0.8	3	868	0.35
2	67	92	72.8	1354	18965	7.1	2	275	0.7	5	4362	0.11
3	142	197	72.1	462	6598	7	58	7784	0.7	2	2520	0.08
4	720	1023	70.4	998	14170	7	18	2550	0.7	6	8450	0.07
5	3795	5404	70.2	3664	52401	7	1012	156771	0.6	8	11605	0.07
6	2581	3694	69.9	1420	20495	6.9	25	3884	0.6	0	191	0
7	5071	7256	69.9	41	615	6.7	24	3858	0.6	0	46	0
8	268	384	69.8	600	9379	6.4	21	4227	0.5	0	144	0
9	559	815	68.6	200	3135	6.4	1	392	0.3			
10	1320	2007	65.8	11	190	5.8	0	10	0			
Total	14612	20982	69.6	8991	128961	7.0	1165	180244	0.6	24	28186	0.09

4.2.4. CD45 depletions are ineffective in large whole blood volumes

Further experiments sought-to clarify the enrichment capability of a single CD45 depletion in whole blood. Unexpectedly, only the control sample with 1,000 spiked PANC-1 cells and no antibody used in the CD45 depletion showed an enrichment effect compared to the negative control of no spiked PANC-1 cells though this did not reach significance (p=0.136) .**Error! Reference source not found.**

4.2.5. CD45 depletion is inconsistent when used in larger volumes of blood from cancer patients

Further clarification of this anomaly was sought using four clinical samples. The patient and tumour characteristics are described in Table 4-2.

Table 4-2. Patient and tumour characteristics of four patients undergoing CD45 depletion of whole blood. The *TP53* variants identified in the corresponding samples are also shown.

Pt No	Diagnosis	Stage		Experimental Set-up			
			Reps	CD45 depletion	No Beads	No Antibody	No depletion
1	PDAC	Unavailable; not resected	8	p.L188R 18% (419/2,383)			
2	PDAC	Unavailable; not resected	7				
3	PDAC	T3,N0,M0	10	p.E271V 32.9% (1126/3423)			
4	PDAC	Unavailable; not resected	6				p.L257P 13% (348/2305)

In Patient 3 the *TP53* variant p.E271V (SIFT class: deleterious) was identified in two out of ten barcodes of the positive control; 32.9% (1126/3423) and 2.88% (82/2848) but not identified in any barcodes of the two negative control samples Figure 4-4. Considered in isolation, this result would be consistent with a CD45 depletion that removes sufficient wild type DNA to enrich CTC derived mutant DNA above the threshold for detection. However, a *TP53* variant was also detected exclusively in the non-depleted whole blood of both patient 4 (p.L257P), the significance of this is unclear**.**

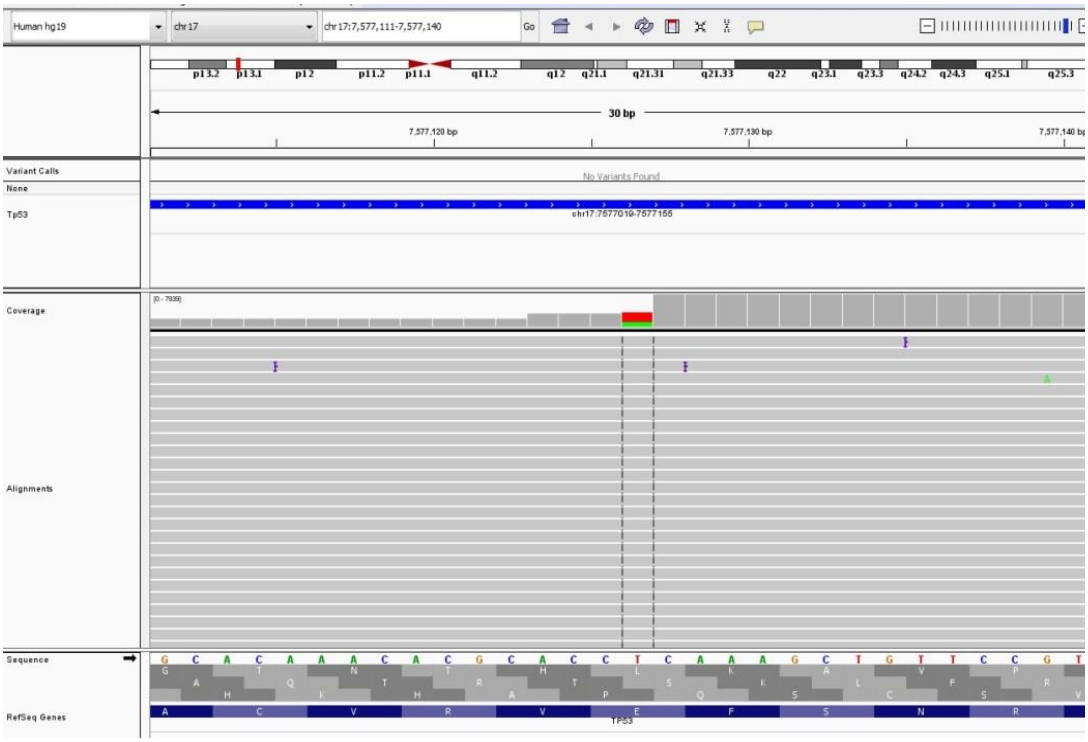


Figure 4-4. IGV screenshot showing p.E271V, c.812A>T variant identified only in CD45 depleted sample.

Next the requirement for the limiting dilution component of the methodology was confirmed. DNA from three experimental samples with low frequency variants were

sequenced, either at a standard concentration of 5ng/µL or diluted to 10 genomes/µL (0.06ng/µL). In all three samples a jackpot effect was seen with elevated variant frequencies observed in the first replicate of samples diluted to 10 genomes Figure 4-5 A-C and Figure 4-6. The fact that the 10 genome samples reached approximately 1% variant frequency, and not 10%, would suggest that perhaps the dilution of samples was not adequate, and more likely represent a dilution of around 100 genomes/µL.

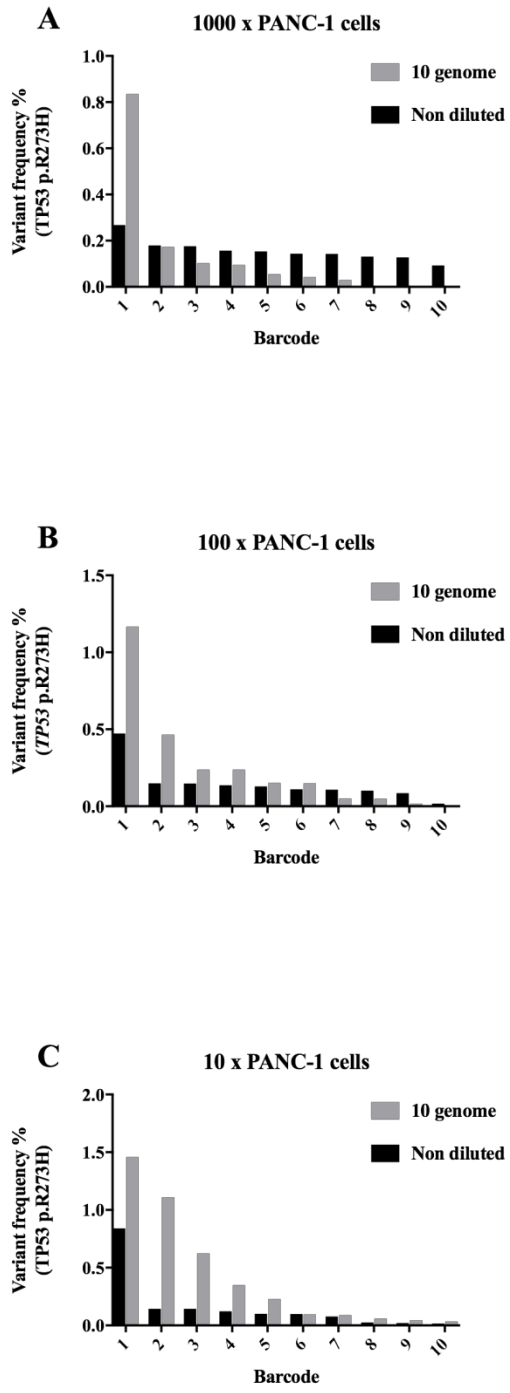


Figure 4-5. The 'Jackpot effect' of limiting dilution demonstrated by the elevated variant frequency seen in barcode 1 of the 10 genome samples. Various numbers of PANC-1 cells: A) 1,000 B) 100 and C) 10 were spiked into 7mL of whole healthy volunteer blood and 1 standard CD45 depletion was performed. The DNA from the enriched sample was either sequenced at a standard dilution of 5ng/μL in 10 replicates (black) or first diluted to 10 genomes/μL (0.06ng/μL) and then sequenced 10 replicates (grey).

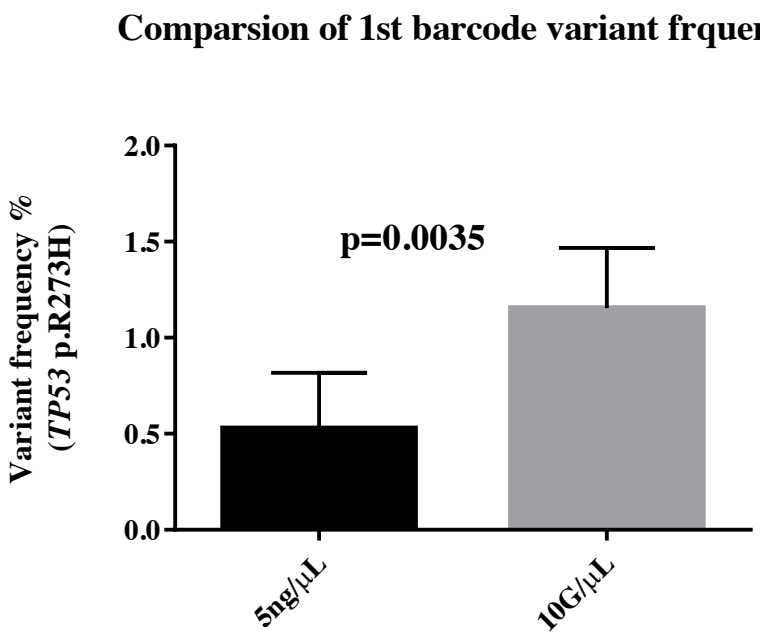


Figure 4-6. Variant frequency is significantly higher at its maximum level when limiting dilution to 10 genome/μL is used compared to standard concentration. Mean frequencies of the three experimental setups Figure 4-5A-C were compared. P value calculated by paired t-test.

4.3 Discussion

In this chapter a method of negative depletion of white blood cells from blood using magnetic ferrous beads covalently bonded to anti CD45 antibodies was developed and assessed. The depletion method showed promise when used in the medium of PBMCs and diluted whole blood by significantly enriching spiked PANC-1 DNA by a demonstrable reduction in wild type DNA (white blood cells). Various combinations of depletions were trialled with the optimal strategy being a single CD45 depletion. Further progress was made by demonstrating improvements in the efficiency of the CD45 depletion by increasing the volume of both the CD45 antibody and the volume of magnetic beads used.

Next, the two components of downstream analysis, limiting dilution and Ion Torrent™ NGS platform were assessed. The limiting dilution methodology was employed to overcome the difficulty of discerning a genuine low frequency variant from an artefact of PCR or sequencing inherent in all sequencing systems. A conservative calculation suggested that 10 replicate samples would be required, and so, 480 primer pairs each with inbuilt barcodes covering the *TP53* gene were designed. The barcodes permit multiplexing of 30 sequencing replicates per sequencing chip, making this a viable approach. With a priori knowledge of the variant (p.R273H) detection down to the 0.1% frequency level was achieved providing a coverage of at least 500 was maintained. Sequencing 10 replicates of standard sample concentration alone was significantly less effective than 10 replicates of 10 genome/μL diluted sample.

The promising results of low volume CD45 depletion were not replicated when the blood volume was scaled up to 7mL of whole blood. A curious anomaly was observed in that the only sample to adequately enrich the PANC-1 signal was the control sample with 1,000 PANC-1 cells and no antibody used in the CD45 depletion. This suggests that at larger volumes the beads are binding to white blood cells in preference of PANC-1 cells and depleting in a manner that is independent of CD45 antibody. Further inconsistencies were seen when clinical samples underwent CD45 depletion, with a *TP53* variant being identified in a non-depleted sample.

The change from small to large clinical blood (7mLs) samples presented resource challenges. The volume of reagents (antibodies and ferrous beads etc) were exponentially greater in the latter. Though only a few patients were trialled the number of samples was much larger as several controls were used for each. A large quantity of

downstream sequencing reagents was also required as each sample required up to 10 serial repeats (10 genome dilutions) to achieve the required threshold of detection. It was pragmatic grounds of resource preservation that a decision was made to move to pre-existing commercial CTC enrichment technologies as alternatives to CD45 depletion.

The use of Ion Torrent NGS and limiting dilution which showed promise here will be continued and used to assess the CTC technologies.

5 ASSESSMENT OF THREE COMMERCIAL CTC ENRICHMENT METHODS

5.1 Introduction

In-house negative depletion strategies for CTC enrichment described in Chapter 2 showed promise in small volumes of blood but were ultimately inconsistent when scaled up to larger volumes. In this chapter three different commercial EpCAM independent CTC enrichment technologies CellSieve™, RosetteSep™ and OncoQuik® which enrich whole blood for CTCs based on size, negative depletion and buoyancy respectively are assessed.

Interrogation of both enriched and non-enriched fractions was performed to investigate the phenotype and genotype of CTCs without *a priori* selection. Other techniques including CTC cell culture and whole blood fractionation were also used.

5.2 Results

5.2.1. CTC enrichment component analysis

The concentration of DNA in both enriched and non-enriched components for all three enrichment methods is shown in Figure 5-1. The CellSieve™ filtration process enrichment process yielded a differential of between 6.39 and 87.7-fold increase in DNA concentration comparing the non-enriched to the enriched fractions Figure 5-1A. The corresponding differentials for Rosette Sep and OncoQuik® were 1.45 to 214-fold (Figure 5-1B) and 18.0 fold (Figure 5-1C) respectively.

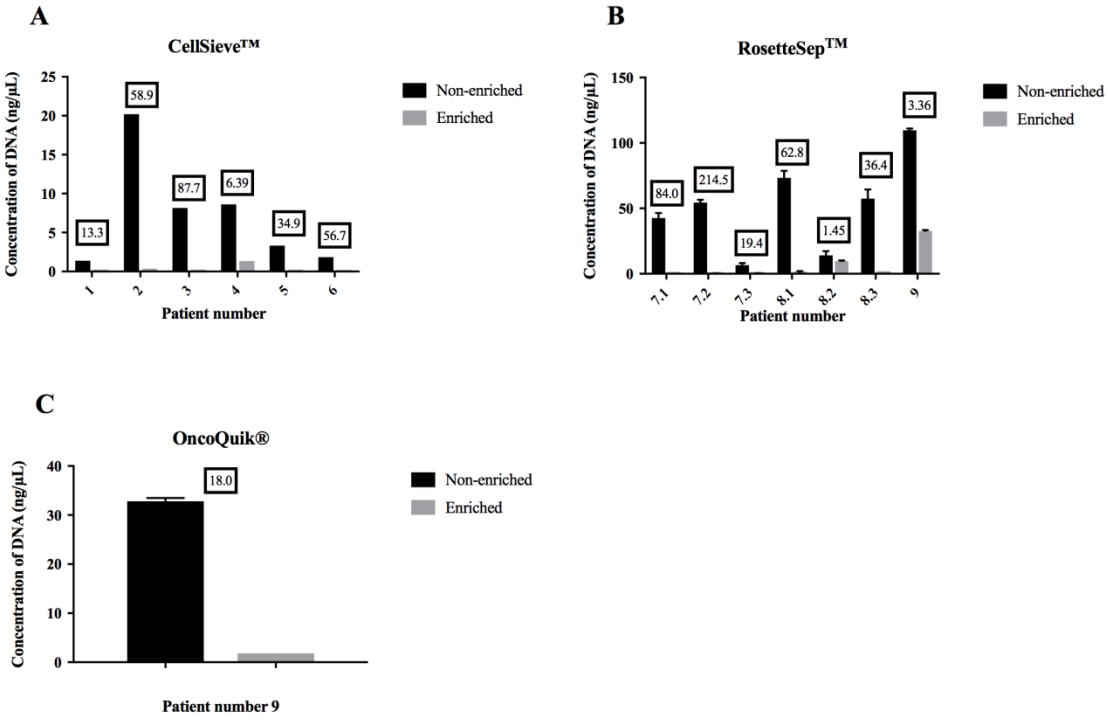


Figure 5-1. Comparison of DNA depletion capacities of three CTC enrichment methods: A) CellSieve™, B) RosetteSep™ and C) OncoQuik®. Patient 7 and 8 were performed in triplicate and Patient 9 was enriched with both RosetteSep™ and OncoQuik®. Error bars representing Mean with SD are shown in B and C as three measurements were taken for each sample. Figures in black boxes represent the factor difference between enriched and non-enriched fractions i.e. concentration of non-enriched/concentration of enriched.

NGS of TP53 was performed on both enriched and non-enriched fractions.

5.2.2. CellSieve™

Non-synonymous variants in the *TP53* gene were identified in 4 out of the 5 patients Table 5-1 and Table 5-2, but in three of these patients the variant was identified in the <10µm non-enriched fraction. Variants were identified in only one of five sequencing replicates in all cases (Table 5-2). *KRAS* mutations were identified in 3 out of 5 patients. All *KRAS* mutations occurred in the enriched > 10µm fraction with the exception of patient 4 who had *KRAS* mutations in both enriched and non-enriched fractions (Table 5-1).

Table 5-1. Mutations in the *KRAS* and *TP53* gene identified on either side of the CellSieve™ filter

	>10µm enriched		**<10µm un-enriched**	
Patient No.	KRAS	TP53	KRAS	TP53
1				p.K132Q
2				
3	p.G12V	p.L257P		
4	p.G12V		p.G12V	p.R282W
	p.G12A			p.K132N
				p.Q167X
				p.K350E
5	p.G12V			p.N288S
				p.E294K

In patient 5 the two *TP53* variants identified in the non-enriched fraction of the CellSieve™ were sought in the corresponding primary FFPE tumour and additional un-enriched whole blood which was sequenced. The first variant, p.N288S was not present in either the primary tumour or the whole blood, however the second *TP53* variant, p.E294K was present at low frequency (0.5%) in the primary tumour, whole blood diluted to 10 genomes/µL (0.3%) and the whole undiluted blood (0.5%) (Table 5-3). The variant p.E294K was looked for in the three cell lines sequenced and identified at a lower level of 2/2,760 (0.07%) in PANC-1, 0/1,166 (0%) in SUIT-2 and 3/2,034 (0.1%) in Mia-Pa-Ca-2.

Table 5-2. Description of all *TP53* variants identified in patients.

Pt No.	Method	Fraction	Protein	Variant	Coverage	%	Effect	SIFT	Somatic
1	CS	Non-enriched	p.K132Q	194	1930	10.1	Missense	Deleterious	19
3	CS	Enriched	p.L257P	2685	7195	37.3	Missense	Deleterious	16
4	CS	Non-enriched	p.R282W	139	848	16.4	Missense	Deleterious	577
	CS	Non-enriched	p.K132N	44	544	8.1	Missense	Deleterious	30
	CS	Non-enriched	p.Q167Ter	96	890	10.8	Nonsense	N/A	43
	CS	Non-enriched	p.K350E	17	156	10.9	Missense	Deleterious	1
5	CS	Non-enriched	p.N288S	254	1938	13.1	Missense	Deleterious	6
	CS	Non-enriched	p.E294K[¥]	161	1143	14.1	Missense	Neutral	5
7	RS-45	Enriched	p.T221Ter	71	387	18.3	Missense	Deleterious	6
	RS-45	Enriched	p.N200S	186	913	20.4	Missense	Deleterious	3
	RS-45	Enriched	p.R175H	260	2255	11.5	Missense	Deleterious	1210
9	OC	Enriched	p.K164R	472	2685	16.1	Missense	Deleterious	2
	RS-45	Enriched	p.E271G	59	343	17.2	Missense	Deleterious	5
10	RS-56	Enriched	p.E271G	46	590	7.7	Missense	Deleterious	5
	RS-36	Enriched		73	2199	3.3			
11	OC	Enriched	p.L265Q	16	129	12.4	Missense	Deleterious	4
	RS-56	Enriched		84	1234	6.8			
	RS-36	Enriched		12	133	9.9			
	OC	Enriched	p.V272M	170	508	33.5	Missense	Deleterious	114
12	OC	Enriched	p.K132Q	194	1231	15.8	Missense	Deleterious	19
13	Cell culture only - N/A								
14	Cell culture only - N/A								

[¥]Variant tracked in patient to whole blood, diluted blood and FFPE tumour: Table 5-3

Table 5-3. Frequency of variant TP53 p.E294K identification in patient 5 in FFPE tumour, whole blood and diluted blood

	FFPE primary tumour			Whole blood			10 genome diluted blood		
Barcode	Variant	Coverage	%	Variant	Coverage	%	Variant	Coverage	%
1	115	33104	0.3	85	25332	0.3	194	998145	0.0
2	126	31391	0.4	2	2020	0.1	12	6113	0.2
3	155	48177	0.3	26	13120	0.2	13	5348	0.2
4	94	31223	0.3	9	5241	0.2	34	20404	0.2
5	71	22159	0.3	1	773	0.1	1	707	0.1
6	16	7522	0.2	2	1029	0.2	11	3479	0.3
7	10	2803	0.4	2	902	0.2	11	2479	0.4
8	28	13244	0.2	0	345	0.0	30	10264	0.3
9	65	18154	0.4	1	1446	0.1	59	23325	0.3
10	1303	264990	0.5	10	1968	0.5	6	2898	0.2

5.2.3. RosetteSep™

Patients 7 to 9 underwent whole blood CTC enrichment with the RosetteSep™ anti-CD 45 kit with both fractions sequenced. Patient 7 was performed in triplicate using three separate anti-CD 45 kits. Three non-synonymous variants were identified in the enriched layer in only one of these triplicates (Table 5-2). No variants were identified in either enriched or non-enriched layer in patient 8. The variant p.E271G was identified in two of the five replicates in the enriched layer of patient 9 (Table 5-2).

5.2.4. OncoQuik®

Patient 9 (above) had whole blood enriched with OncoQuik® as well as Rosette Sep in the same experiment. Enriched and non-enriched fractions were sequenced in both instances. Non-synonymous variants were identified exclusively in the enriched fractions using both methods however the variants were different. The OncoQuik®

sample identified a p.K164R variant in one of five replicates and the RosetteSep^{TM} identified the p.E271G variant described above Table 5-2.

The diagnosis of the 14 patients recruited for this chapter are shown in Table 5-4. All 14 patients were recruited and sampled on the ward in the Royal Liverpool NHS Trust in accordance with the study protocol described in the methods section. The timing of sampling in relation to their surgery and treatment is also shown in Table 5-4. Patient and tumour characteristics of 14 patients recruited to study

Table 5-4. Patient and tumour characteristics of 14 patients recruited to study

Pt No.	Enrichment	Diagnosis	Resected	Post surgery (days)
1	CellSieve™ and Cell culture	PDAC	Yes	3
2	CellSieve™ and Cell culture	PDAC	Yes	5
3	CellSieve™ and Cell culture	IPMN tumour	Yes	9
4	CellSieve™	NET pancreas	Yes	3
5	CellSieve™	Ampullary adenocarcinoma	Yes	14
6	CellSieve™	Metastatic PDAC	Yes	3
7	Rosette Sep™	Metastatic gastric adenocarcinoma	No, palliative bypass	6 months
8	Rosette Sep™	Locally advanced PDAC	No, palliative bypass	6 months
9	Rosette Sep™, OncoQuik®	Locally advanced PNET	No, locally advanced	
10	Rosette Sep™, OncoQuik®	PDAC	Yes	8
11	Rosette Sep™, OncoQuik®	PNET	Yes	10
12	OncoQuik® and Fractionation	PDAC	No, locally advanced	
13	Cell culture	Intra-hepatic cholangio	No, locally advanced	
14	Cell culture	PDAC	No, locally advanced	

5.2.5. OncoQuik® and RosetteSep™ comparison

Since variants were exclusively identified in the enriched fractions of RosetteSep™ and OncoQuik® but not CellSieve™ further direct comparisons were performed analysing just the enriched fractions for RosetteSep™ and OncoQuik®. Patients 10 and 11each had OncoQuik®, RosetteSep™ anti-CD 36, and RosetteSep™ anti-CD 56 enrichment performed.

In patient 10 the variant p.E271G was identified in 4 out of 5 replicates in the RS-56 enrichment (ranging from 1.8 to 7.8%). This variant could also be found in 3 out of 5 replicates in RS-36 but at lower frequencies (range: 0.7 to 3.3%). There was no evidence of this variant in the OncoQuik® enrichment Table 5-2.

In patient 11 however the variant p.L265Q was detected in all 5 replicates of the OncoQuik® enrichment (range: 5.5 to 12.4%), in two of five replicates in RS-56 (6.8 and 4.8%) and in one of five replicates in RS-36 (9.0%) Table 5-2. Patient 11 also had the variant p.V272M exclusively detected in one of five replicates in the OncoQuik® enrichment but was not found in either RosetteSep™ enrichment.

These variants identified in the OncoQuik® vs RosetteSep™ comparison were taken forward for confirmation using mutation specific PCR. Both p.L265Q Figure 5-2A, which was present in all three enrichment methods of patient 11 and p.V272M Figure 5-2B variants were confirmed using mutation specific PCR in the OncoQuik® enriched sample from patient 11, though p.E271G could not be reproduced with the designed primers in patient 10.

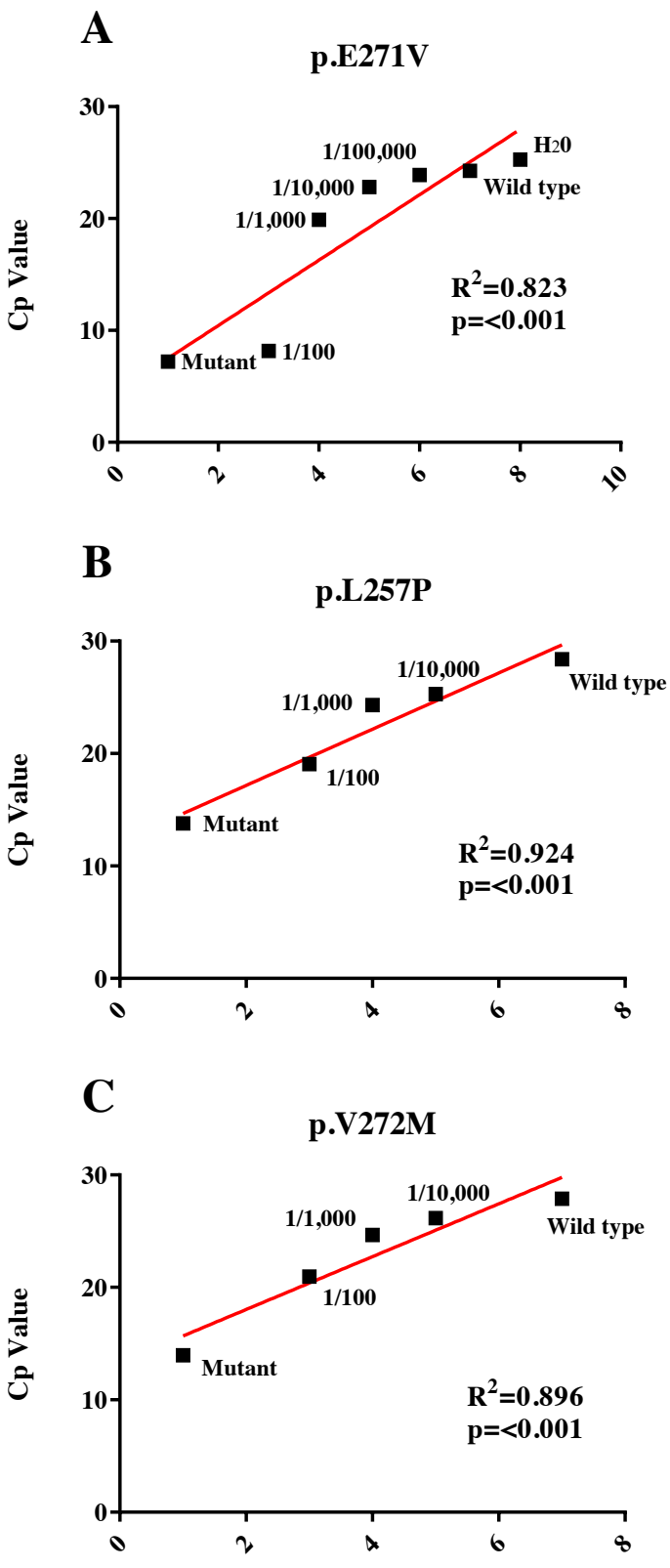


Figure 5-2. Relationship shows good specificity of primers designed for three variants: A) p.E271V, B) p.L257 and C) p.V272M.

5.2.6. Fractionation experiments

Patient 12 underwent OncoQuik® enrichment and fractionation of whole blood into 21 aliquots both pre and post operatively. The pre-operative OncoQuik® enrichment and 21 fractions did not identify any variants. The post-operative sample (unresectable PDAC so tumour remained *in situ*) identified the variant p.K132Q in the OncoQuik® enrichment in 3 of 9 replicates at the following frequencies: 194/1231 (15.8%), 130/1132 (11.5%) and 342/3305 (10.3%). The same variant p.K132Q was also detected in aliquot numbers 1 (132/1189, 11.1%), 2 (47/365, 12.9%), 3 (67/108, 62.0%) and 10 (63/602, 10.5%) Figure 5-3.

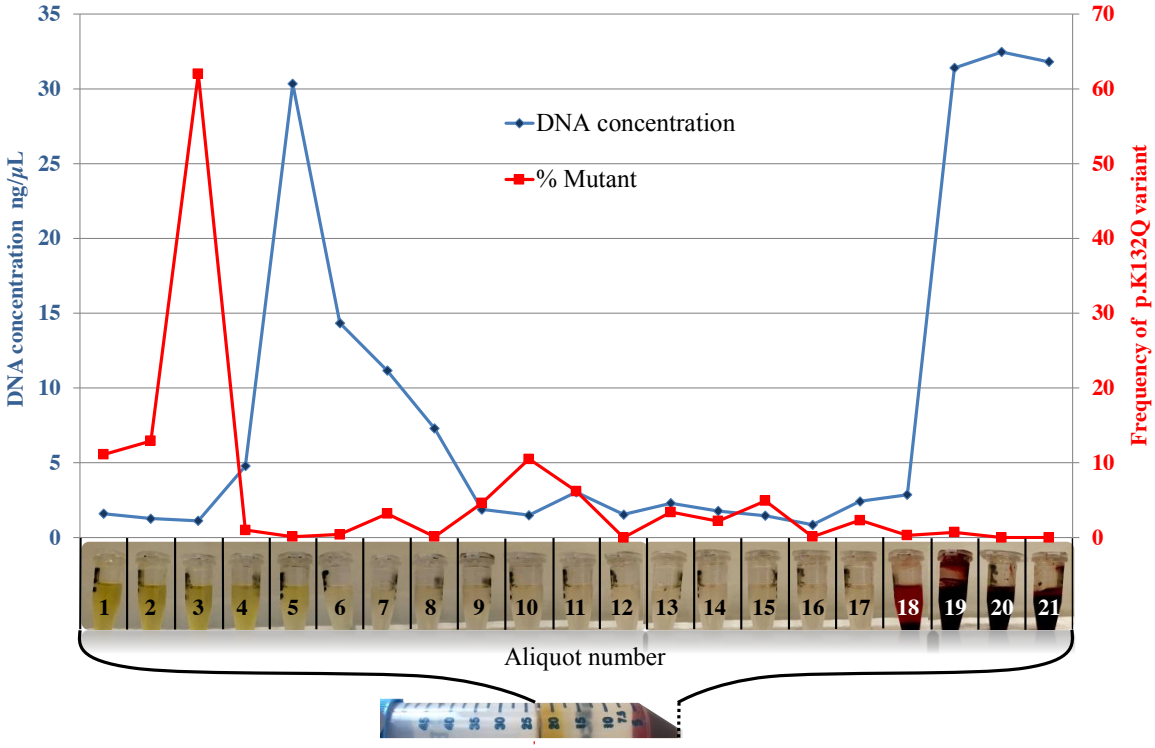


Figure 5-3. Graph showing frequency of p.K132Q variant and concentration of DNA in each of the 21 aliquots from patient 10

5.2.7. Downstream NGS of Cell Culture

Cells isolated from patients 1,2,3,13 and 14 underwent cell culture by Dr. Nick Bryan. Exons 5,6,7 and 8 of the TP53 gene were successfully amplified by PCR from the 5 samples Figure 5-4. All samples were wild type for both *KRAS* and *TP53* with clean sequence seen on Sanger sequencing Figure 5-5.

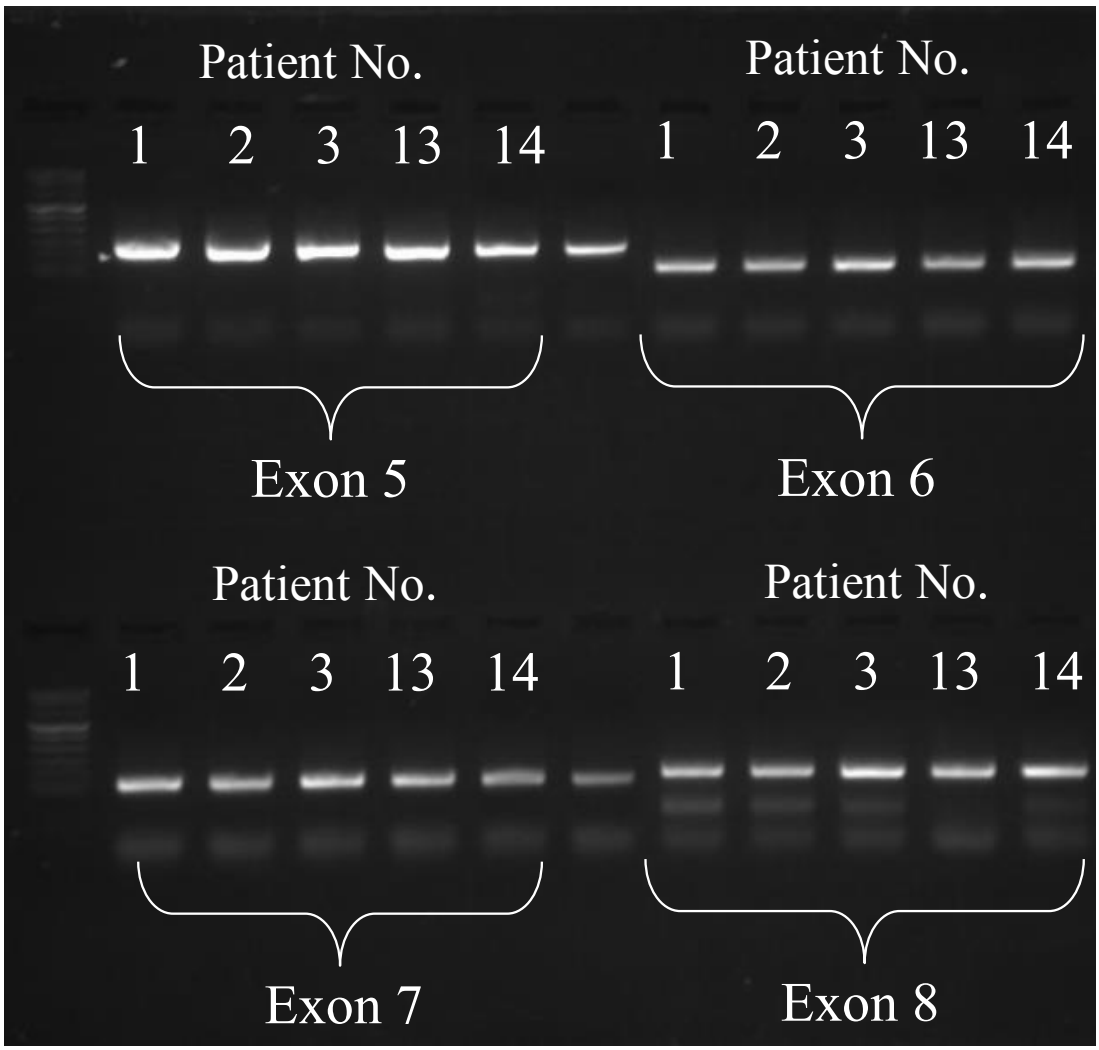


Figure 5-4. 1.8% agarose gel showing bands at expected positions for Exons 5, 6, 7 and 8 of the TP53.

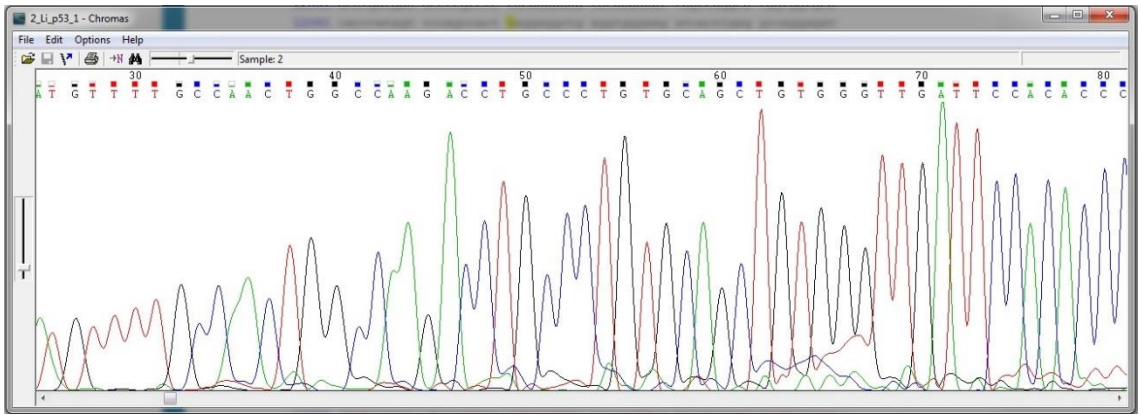


Figure 5-5. Screenshot of Sanger sequencing analysis from patient 1 showing section of Exon 5 of TP53 gene using software Chromas version 2.4.

5.2.8. TP53 NGS development

Sequencing an identical library preparation consisting of 30 barcoded TP53 amplicons on another date and chip, showed 100% concordance with variant identification. The mean difference in variant frequency between the first and second run was 3.09%, 95% CI [1.67, 4.51] Table 5-5.

Table 5-5 Comparison of the variant reads of the same library on two separate dates

Date run	Patient ID	Variant	Variant reads	Total reads	% variant	Δ%
03/12/2014	4	p.K132N	42	438	9.59	1.49
17/12/2014			44	544	8.10	
03/12/2014	4	p.Q167X	106	767	13.82	3.03
17/12/2014			96	890	10.79	
03/12/2014	4	p.K305E	21	304	6.91	3.99
17/12/2014			17	156	10.90	
03/12/2014	4	p.R282W	128	815	15.71	0.69
17/12/2014			139	848	16.39	
03/12/2014	5	p.N288S	254	1938	13.10	3.77
17/12/2014			196	1162	16.87	
03/12/2014	5	p.E294K	372	1891	19.67	5.57
17/12/2014			161	1143	14.10	

5.3 Discussion

In this chapter three commercial CTC enrichment products were investigated. It is acknowledged that there is heterogeneity with respect to both histology and peri-operative timing of sample collection in some of the patients. However, in this chapter the aim was to select the most appropriate technologies for a further and more detailed assessment, in a more homogeneous patient population in chapter six. Limited resources meant that it would not have been possible to interrogate all three technologies in the larger patient population and hence a degree of heterogeneity was accepted here. With regards to histology, patient samples were collected and analysed in the peri-operative period, in some instances, before the final histology results were available. As samples were required to be processed fresh (within 4 hours) it was not possible to know the histology result at the time of analysis; therefore a pragmatic approach was taken to include these samples in the assessment of CTC technology despite being strictly non-PDAC. As all these different histological cancer types are known to have CTCs which behave in a similar fashion they represent a suitable sample to assess the technology, though, of course, PDAC would have been preferable. In the next chapter where whole patient cohorts are assessed, and the need for homogeneity is greater, such patients were excluded.

With regards to the peri-operative timing of sample collection, again, a pragmatic approach was necessary. Sample processing could only be performed at specific times of the week as it required the additional expertise of the technician Mr Neal Rimmer and Clinical engineer Dr. Nick Bryan until the methods were established and I could perform the enrichment independently. Whilst sample analysis was scheduled for the morning of surgery, inevitable cancelations of surgical lists associated with real life clinical practice meant that, on some occasions, post-operative patients available on the ward had to be sampled instead.

The exact relationship between CTCs and surgery has not been clearly defined. Some studies have shown that manipulation of the tumour during surgical resection significantly increases the number of CTCs detectable in the blood (S. Ito et al., 2002)(Park et al., 2012). An increase in CTCs after surgery has also been associated with early recurrence of gastric cancer(Q. Zhang et al., 2018) and another recent prospective trial demonstrated that CTCs detected post-operatively, but not those

detected pre-operatively, were associated with a higher risk of recurrence(Yang et al., 2018). Whilst the clinical question of the presence of CTCs and their correlation with recurrence is a valid one, it was not the purpose of this chapter to investigate it, nor would it have been possible to do so with such small numbers. As previously stated, the aim was to select the most appropriate CTC enrichment technology for a further and more detailed assessment in chapter six. The available research suggests that both pre-operative and post-operative CTC sampling is valid in this setting and provides justification for doing so in this chapter. Homogenous sampling would have been preferable, but for the logistical reasons discussed was not possible in this chapter. In chapter six the timing of sampling in the peri-operative period is standardised throughout the cohort.

Patient blood samples produced two fractions from whole blood, generating a differential in DNA concentration but by different mechanisms. Sequencing both of these fractions gave additional information upon which to judge its performance. CellSieve™ appeared to give inconsistent results with variants being identified in the non-enriched fractions in 3 out of 6 patients. This could perhaps be due to the increasingly recognised presence of mesenchymal sub-populations of CTCs(Kalluri & Weinberg, 2009) which are smaller than the conventional epithelial CTCs which the filter is designed for(D. L. Adams et al., 2014) and these mesenchymal cells might have been able to pass through the 10μm pores of the filter. However, the non-enriched fraction contains the WBCs and one would expect the wild type signal from these to mask any variant signal from mesenchymal CTCs, unless CTCs are present in much higher quantities than previously thought(Gorges et al., 2012). Evidence that the variants identified in the non-enriched (< 10μm portion) are tumour derived was provided by identifying the identical non-synonymous *TP53* variant in the primary tumour. Yachida et al showed that a 64% of mutations were “founder” mutations (present in primary and metastasis) whereas “progressor” mutations, which evolved after leaving the primary tumour, and therefore not identified in the primary, were less frequent (36%)(Shinichi Yachida et al., 2011). The very low frequency of variant seen in whole blood also confirms that it is not currently possible to identify variants in blood without a priori knowledge of the variant and an enrichment step.

The OncoQuik® enrichment system is based on the assumption that CTCs are more buoyant than WBCs and RBCs(Gertler et al., 2003a). The results of the fractionation experiment were consistent with a population of less buoyant CTCs. The variant frequency peak was identified in the 3rd fraction out of 21, just above the 4th fraction corresponding to the PBMC layer (shown by a peak in DNA concentration). In this experiment no variants were identified on the pre-operative sample, but were identified on the first post-operative day where the tumour was dissected but not removed. Studies have shown that intra-operative tumour handling increases the quantity of CTCs identified intra and post-operatively, perhaps by the dislodgment of tumour cells from the primary tumour(Papavasiliou, Fisher, Kuhn, Nemunaitis, & Lamont, 2010)(Pesta, Fichtl, Kulda, Topolcan, & Treska, 2013)(Sawabata, Funaki, Hyakutake, & Shintani, 2016)(Juratli et al., 2014). Though this was only investigated in a single patient the outcome is consistent with this theory.

RosetteSep™ and OncoQuik® showed promise yielding cells with *TP53* variants in enriched fractions, though such variants were not seen in every sample, and could not be replicated in some experimental repeats. Head to head comparisons showed different variants were identified within the same patient. RosetteSep™ employs a purely negative selection approach whilst OncoQuik® enriches on the basis of CTC density. It has been experimentally demonstrated that CTCs are heterogeneous in both genotype and phenotype. It is feasible that part of this heterogeneity are separate populations associated with different *TP53* mutations, each population having different physical characteristics, some more likely to be enriched by RosetteSep™ and some more likely to be enriched by OncoQuik®.

Consistency was demonstrated in patient 11 with the same non-synonymous *TP53* deleterious variant being identified in two different RosetteSep™ enrichments and OncoQuik® which was subsequently confirmed with mutation specific PCR. Consistency was also seen in the Ion torrent™ calling software, with 100% concordance for variants and broadly similar frequencies.

The elusive goal of culturing CTCs from whole blood was not realised in this chapter. A population of cells which do not appear to be haemopoetic in origin grew out but were excluded as CTCs on the basis that no mutations in *TP53* or *KRAS* were identified. Restraints on laboratory and personnel availability combined with the requirement to process all clinical samples within 4 hours limited patient samples to surgical inpatients

with Hepatobiliary type cancer. Consequently, the cohort of patients in this chapter are a little heterogeneous.

Given these promising result both RosetteSep™ and OncoQuik® were taken forward to investigate a larger more homogenous cancer cohort with appropriate controls in the next chapter.

6 APPLICATION OF CTC TECHNOLOGIES TO PDAC COHORTS

6.1 Introduction

In Chapter 5, three methods of CTC enrichment were assessed; Rosette Sep™ and OncoQuik® performed best and were taken forward to examine CTCs in two cohorts of patients with PDAC. Comparison with primary and metastatic tumour, as well as serial sampling was performed. The two main aims of this chapter were firstly to determine if variants could be tracked across multiple samples of the same patient and secondly, to determine if clonal sweeps (cycles of chemotherapy) affected the mutational profile of CTCs.

6.2 Results

6.2.1 The Oncomine™ solid tumour DNA panel identifies low frequency mutations present in PDAC.

The three cell lines used are known to have mutations in both *TP53* and *KRAS* genes(Moore et al., 2001) outlined in Table 6-1.

Table 6-1. Mutations present in the three cell lines used.

Cell line	KRAS	TP53
PANC-1	p.G12D	p.R273H
Mia-Pa-Ca-2	p.G12C	p.R248W
SUIT-2	p.G12D	p.R273H

The Oncomine™ Solid Tumour DNA panel identified all the mutations present in Mia-Pa-Ca-2, SUIT-2 and PANC-1 in addition to 21 other genes commonly mutated in PDAC **Error! Reference source not found.**. The coverage was such that mutations could be consistently identified without drop-out down to a frequency of 1% (Table 6-2 and Figure 6-1).

Table 6-2 Expected versus observed mutation in MIA-PA-CA-2 cell line

Expected %	***TP53* p.R248W**			***KRAS* p.G12C**		
	Variant	Total	%	Variant	Total	%
100	2999	3072	97.6	4250	4285	99.2
99.99	257	261	98.5	379	379	100
	1819	1859	97.8	2430	2445	99.4
99.9	334	343	97.4	605	610	99.2
	1050	1082	97	1146	1152	99.5
99	418	437	95.7	428	437	97.9
	770	802	96	802	821	97.7
90	249	290	85.9	317	354	89.5
	235	268	87.7	262	330	79.4
0	0	1053	0	1	1405	0.1
	35	3214	1.1	3	6146	0

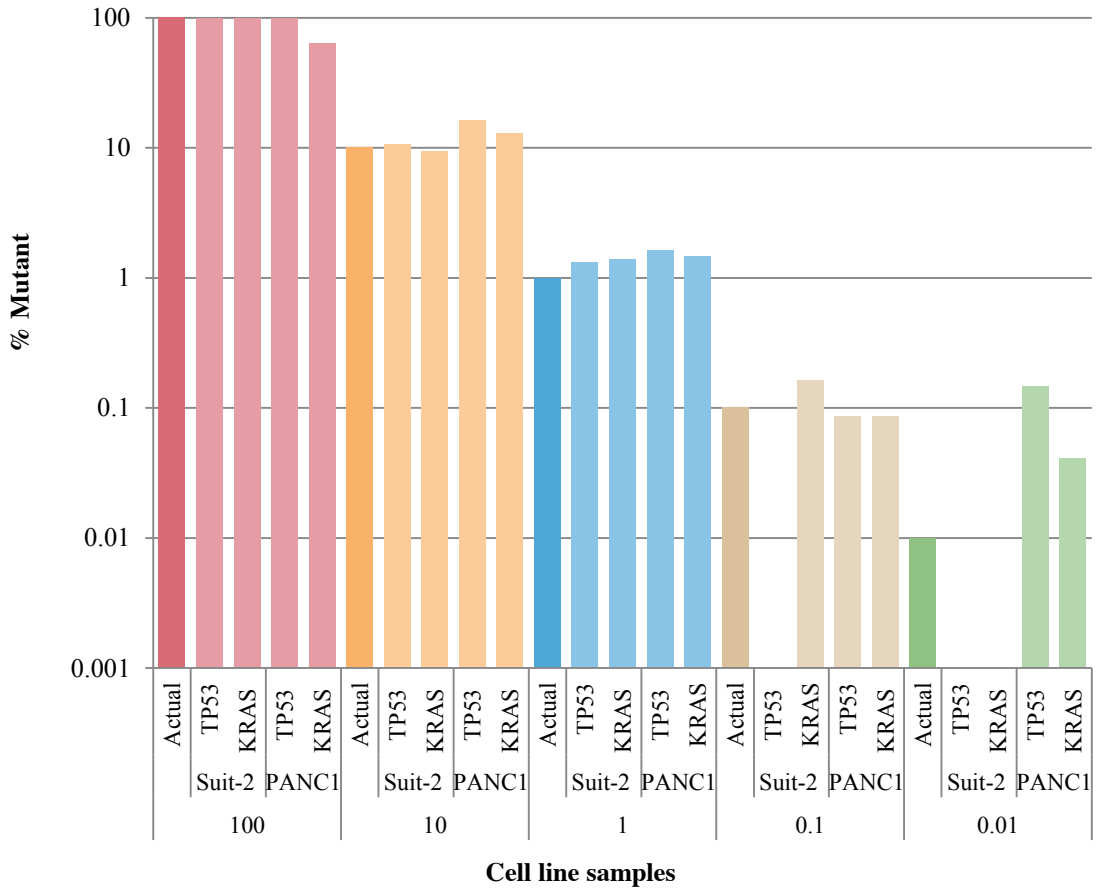


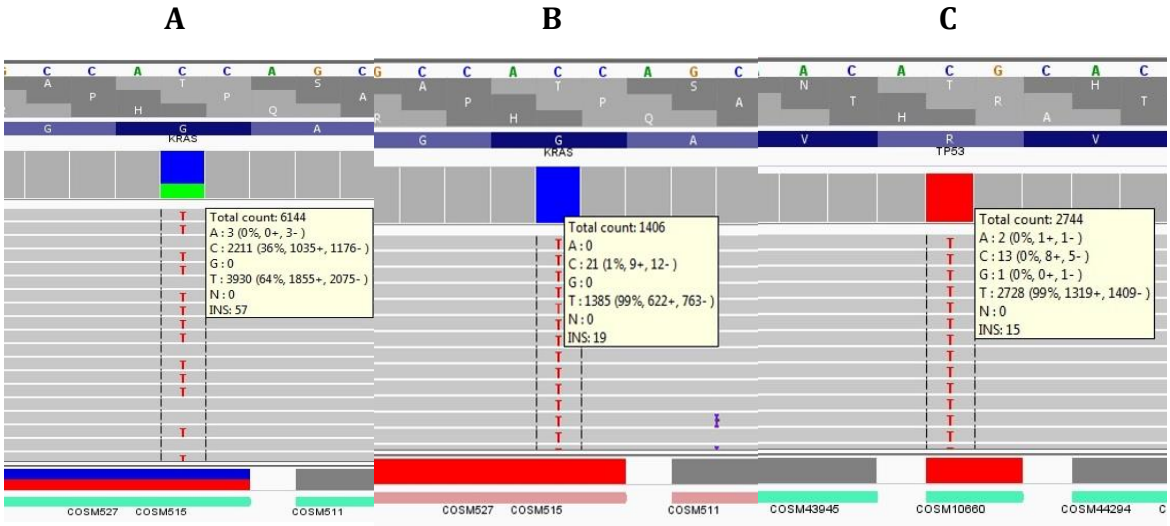
Figure 6-1. Identification of known mutations present in SUIT-2 and PAC-1 cell lines using the Oncomine™ Solid tumour DNA panel. X-axis is logarithmic scale

With the exception of *KRAS* p.G12D in the pure PANC-1 sample Figure 6-2A, all mutations were identified at broadly the expected frequency. In this example, 3,930 variant reads were called out of a total of 6140 (64%)**.**

Figure 6-2A. In the pure SUIT-2 sample the same *KRAS* mutation was called very close to 100% with 1385 variant calls out of a total of 1,406 reads (98.5%)

Figure 6-2B. Equally, the *TP53* mutation p.R273H in the pure PANC-1 was sample was also very close to 100% with 2728 calls out of a total of 2,744 reads (99.4%)

Figure 6-2C. Of note, in chapter 4 Table 4-1 the TP53 sequencing data revealed only 70% mutant sequence. For this section a different batch was used which would explain the discrepancy observed.



Pure PANC-1 sample.	Pure SUIT-2 sample.	Pure PANC-1 sample.
KRAS p.G12D	*KRAS* p.G12D	*TP53* p.R273H

Figure 6-2. IGV screenshots of mutations identified in cell lines. A) Pure PANC-1 sample showing KRAS p.G12D at 65%, B) Pure SUIT-2 sample showing KRAS p.G12D at 99% and C) Pure PANC-1 sample showing TP53 p.R273H at 99%.

6.2.2 Patient characteristics of cancer and control patients in OncoQuik® cohort

Overall 35 patients who were undergoing pancreatic resection were selected for the OncoQuik® cohort. All these patients had blood taken pre-operatively on the morning of surgery which was enriched using OncoQuik® and subsequently sequenced as described. Once the histology was available in the post-operative period 6 patients were excluded as histological examination did not confirm PDAC. The flow diagram describes which patients were excluded from the study with reasons Figure 6-3.

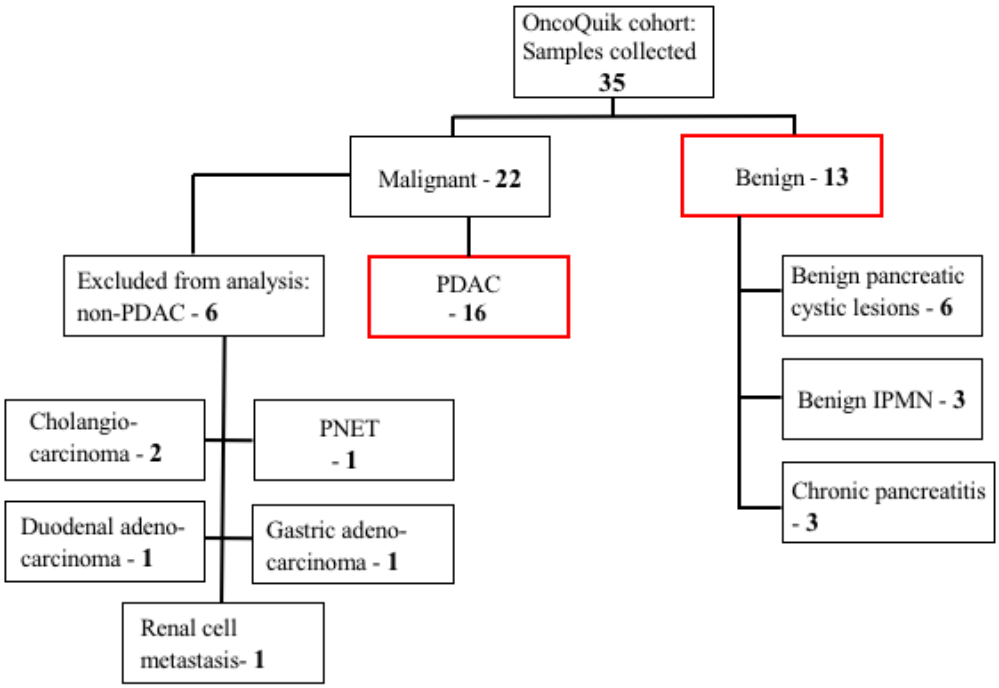


Figure 6-3 Flow diagram of samples collected for the OncoQuik® cohort

The OncoQuik® cohort consists of 16 patients in the cancer group and 13 patients in the control group. Patient characteristics are described in Table 6-3.

Table 6-3. Characteristics of cancer and control patients in the OncoQuik® cohort

		Cancer	Control	p
		RosetteSep™ (n=15)	(n=10)	
Female		7	5	1
Median age years (range)		69 (55-84)	67.5 (41-89)	0.3848
Histology	PDAC	15	n/a	
Resected	Yes	7		
	No	8		
Grade	Well	2		
	Moderate	0		
	Moderate-poor	3		
	Poor	1		
	Unknown	1		
T stage*	T1	0		
	T2	1		
	T3	8		
	T4	6		
N stage*	N0	6		
	N1	9		
	Not reported	0		
M stage*	M0	14		
	M1	1		
R status	R0	4		
	R1	0		
	Unknown	11		

* The number of non-diabetics will include a number of Type 3c diabetic patients who were undiagnosed at time of diagnosis with cancer. Up to 60% of patients with PDAC are estimated to have type 3c diabetes (Chari, et al., 2008). P values determined by Mann Whitney U, Pearson's chi-squared or Fisher's Exact tests as appropriate. The tumour characteristics of the OncoQuik® cancer group are described in Table 6-4.

Table 6-4. Tumour characteristics of cancer patients in OncoQuik® and Rosette Sep™ cohort

		OncoQuik® (n=16)	RosetteSep™ (n=15)
Histology	PDAC	16	13
Resected	Yes	11	7
	No	5	8
Grade	Well	0	2
	Moderate	6	0
	Moderate-poor	2	3
	Poor	3	1
	Not reported/Unknown	0	1
	Un-resected	5	8
T stage	T1	2	0
	T2	0	1
	T3	10	8
	T4	4	6
N stage	N0	4	6
	N1	11	9
	Not reported	0	0
M stage	M0	11	14
	M1	5	1
R status	R0	2	4
	R1	8	0
	Unknown	1	3
	Un-resected	0	8

6.2.3 Identification of mutations in Cancer and Control groups

Variants with a frequency greater than 5% were-identified in 3/16 (18.8%) cancer patients. Mutations were only identified in the cancer group, but with such small numbers this did not reach significance (P=0.162).Table 6-5

Table 6-5. Comparison of number of patients with variants >5% in the OncoQuik® cohort. P=0.162 using Fisher's exact.

	>5% mutation	No mutation >5 %	Total
Cancer	3	13	16
Control	0	13	13
Total	5	30	35

When the 5% frequency threshold was removed, variants were identified in 6/16 (40%) cancer patients compared with 2/13 (15.4%) control patients showing a trend towards significance (P=0.18) (Table 6-6 and Figure 6-4).

Table 6-6. Comparison of variants at ANY frequency in cancer and control patients in the OncoQuik® cohort. P=0.07 using Fiseher's exact

	Any mutation	No mutation	Total
Cancer	6	10	16
Control	2	11	13
Total	8	21	29

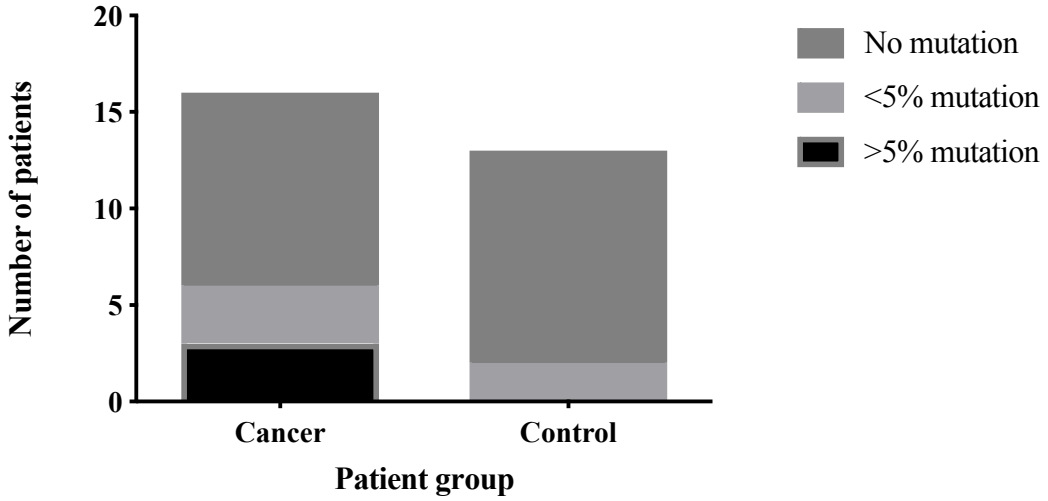


Figure 6-4. Comparison of mutation frequency identified between cancer and control groups in the OncoQuik® cohort

No mutations were identified in both replicates of the same patient. A description of all the mutations over the 5% threshold from Table 6-5 is shown in Table 6-7.

Table 6-7 Descriptions of mutations over the 5% threshold in the OncoQuik® cohort

Patient number	Gene	DNA	Protein	Variant	Total	%
3	STK11	c.817G>A	p.Ala273Thr	6	98	6.1
4	ERBB2	c.2543C>T	p.Ala848Val	80	1309	6.1
	CSDE1	c.82T>C	p.Phe28Leu	32	587	5.5
	TP53	c.1001G>A	p.Gly334Glu	46	607	7.6
10	ALK	c.3604G>A	p.Gly1202Arg	78	1465	5.3
	TP53	c.997C>T	p.Arg333Cys	60	1009	5.9
	SMAD4	c.1529G>A	p.Gly510Glu	61	304	20.1
	STK11	c.1043A>G	p.Asp348Gly	32	512	6.3
	DDR2	c.298G>A	p.Val100Met	92	646	14.2

6.2.4 Serial sampling of OncoQuik® cohort shows mutations no longer present after resection in cancer patients.

Only one cancer patient had repeat blood sampling and OncoQuik® enrichment performed in the follow-up period after resection Table 6-8. Three patients in the control group had repeat blood sampling after resection for benign disease

Table 6-9. In the cancer patient (patient 7), no variant was identified either in the pre or post-operative sample, however this patient had a borderline resectable tumour and received pre-operative chemotherapy in an attempt to render the tumour resectable. Subsequent histological examination of the resected pancreas revealed only a 'tiny' focus of PDAC remaining. In these patients it is possible to speculate the proportion of cancer of the total in the sample. For example in patient number 5 Table 6-11 a total of 8% cancer cells is estimated, with a homozygous mutation in TP53 (7.6%) and heterozygous mutation in MAP2K1 gene (3.8%). Again, in patient 10 Table 6-11 one could speculate a total of 9% cancer cells with heterozygous TP53 (4.3%) and homozygous FGFR3 (9.7%).

In the control group, none of the three patients had variants identified pre-operatively. In one patient however, after resection of a duodenal gastro-intestinal stromal tumour (GIST), a KRAS mutation variant p.Leu23Gln was identified at 30/346 (8.7%).

In all patients with serial sampling, polymorphisms could be tracked across patients to confirm sample identity.

Table 6-8. Mutation identified in blood using OncoQuik® enrichment and NGS with Oncomine™ panel before and after surgery in cancer patient

			Before resection	Interval (days)	After resection
Diagnosis	**Pt No.**	**Stage**	**Mutation**		**Mutation**
PDAC	7	T1 N0 M0	None	31	None

Table 6-9. Comparison of mutations identified in blood using OncoQuik® depletion and NGS with Oncomine™ panel before and after surgery in control patients

	Before resection	Interval (days)	After resection				
Diagnosis	**Mutation**		**Mutation identified**				
			Gene	Protein description	Variant	Total	%
Duodenal GIST	None	121	KRAS	p.Leu23Gln	30	346	8.7
IPMN	None	76	None				
Duodenal GIST	None	144	None				

6.2.5. Sequencing of primary tumour and metastasis in the OncoQuik® cohort

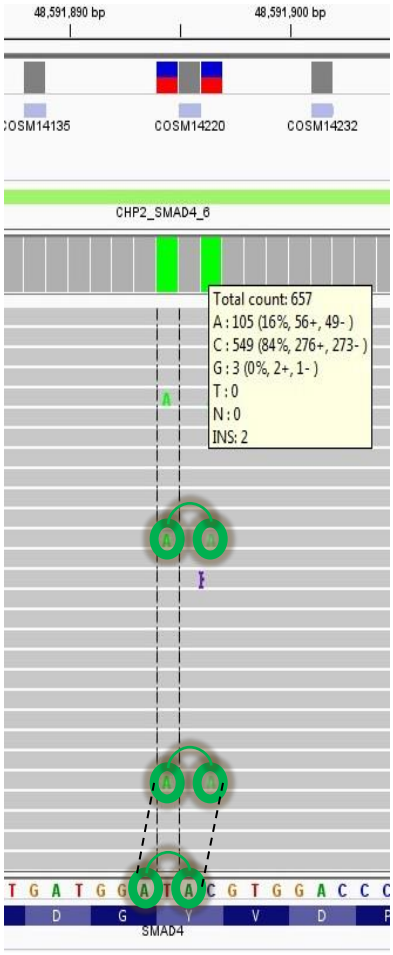
Variants of any frequency were identified in 6 patients in the OncoQuik® cohort. FFPE tumour samples were available for analysis in 5 of these. A metastatic peritoneal nodule was sequenced in patient 4 who had un-resectable disease. The primary tumour was analysed in 4 patients and all had nodal disease (N1), in these patients the involved lymph node was also identified and sequenced. Comparisons of the mutations found in the solid tumour were made to that of mutations found in the corresponding OncoQuik® enriched blood and vice versa. None of the mutations identified in the solid tumour were identified in the corresponding OncoQuik® enriched blood samples Table 6-11. Equally, none of the variants identified in the OncoQuik® enriched samples were seen in the corresponding solid tumour Table 6-12. Clearly, these results show very poor correlation between primary tumour and OncoQuik® enriched samples of the peripheral blood. The solid tumour samples are derived from FFPE of the resected specimen. As described in the methods chapter, attempts were made to maximize the percentage of

tumour DNA within the sample by scalpel dissection of the block guided by histopathological targeting. It is possible to estimate the percentage tumour cells in the sample by examining the sequence results by taking into account possible filed cancerization and heterozygous and homozygous mutations.

Primary tumour and metastatic lymph node pairs were available for comparison in three patients (patient 2, 12 and 17)**,** as the sequencing for the lymph node sample failed in two patients (patient 5 and 22). In all three primary versus lymph node comparisons, the same mutations identified in the primary tumour were also present in the lymph node. In patient 2 however an additional mutation was superimposed in the lymph node (TP53 p.Arg196Ter) which was not present in the primary tumour. This additional mutation could represent a branched evolution facilitating the step to lymph metastasis.

In patient 17 two additional variants in the SMAD4 gene (p.Tyr353Asn and p.Tyr353Ter) were identified in the primary tumour but not in the lymph node. However, on examining these variants in the integrative genomic viewer they appear to be due to an alignment errors. The two variants are both caused by two ‘A’ nucleotides separated by a T nucleotide, a shift of the reference genome by one nucleotide.

Figure 6-5. Intergrative genomic viewer (IGV) screenshot of two false positive mutations in *SMAD4* gene of primary tumour of patient 2.



In all patients with multiple samples, linking polymorphism were sought and identified, for example Table 6-10.

Table 6-10 Example of linking polymorphism (heterozygous for p.Glu168Asp in MET gene) across all biospecimens in Patient 17.

	Variant	Coverage	%
Blood rep 1	362	811	44.6
Blood rep 2	526	1079	48.7
Primary	511	1157	44.2
Lymph node	455	1344	33.6

Table 6-11. A comparison of mutations identified in solid tumour material with the corresponding OncoQuik® enriched blood samples

	Mutation		Primary			Lymph node			Metastatic			Estimated tumour %	Blood 1st rep			Blood 2nd rep		
Pt No.	Gene	Protein	Variant	Total	%	Variant	Total	%	Variant	Total	%		Variant	Total	%	Variant	Total	%
3	TP53	p.Val157Gly							22	180	12	12	0	2282	0	0	3999	0
	KRAS	p.Gly12Asp							59	657	9		1	3085	0	0	3287	0
	SMAD4	p.Glu167Ter							18	518	3.5		0	3008	0	0	3970	0
4	TP53	p.Arg213Ter	146	742	20	Failed						20	0	799	0	0	438	0
	KRAS	p.Gly12Asp	405	2566	16								0	1540	0	0	889	0
10	TP53	p.Arg175His	201	655	31	281	3995	7				7	0	1243	0	0	1186	0
	KRAS	p.Gly.12Asp	1657	3984	42	25	559	4.5					0	704	0	0	1845	0
13	KRAS	p.Gly12Ser	945	2506	38	291	1276	23				40	0	1100	0	0	841	0
	SMAD4	p.Gly510Ala	271	664	41	67	706	9.5					0	1265	0	0	1027	0
		p.Tyr353Asn	106	653	16	See Figure 6-6							0	934	0	0	1224	0
		p.Tyr353Ter	107	656	16								0	934	0	0	1224	0
14	TP53	p.Cys275Arg	20	228	8.8	Failed						8	0	1221	0	0	1221	0
	KRAS	p.Gly12Val	81	834	9.7								0	1819	0	0	1819	0

Table 6-12 A comparison of mutations found in OncoQuik® enriched blood with the corresponding solid tumour material

Pt No	Mutation		Blood			Primary			Lymph node			Metastatic		
	Gene	Protein	Variant	total	%	Variant	total	%	Variant	total	%	Variant	total	%
3	TP53	p.His193Arg	49	2046	2.4							0	190	0
		p.Thr81Ile	41	1528	2.7							0	102	0
	SMAD4	p.Val409Ile	140	3821	3.7									
	STK11	p.Ala273Thr	6	98	6.1							0	349	0
	PTEN	p.Ile5Val	71	1648	4.3							0	379	0
	MET	p.Ala991Val	148	4000	3.7							0	77	0
4	TP53	p.Gly334Glu	46	607	7.6									
	ERBB2	p.Ala848Val	80	1309	6.1									
	CSDE1	p.Phe28Leu	32	587	5.5									
	FGFR1	p.Ser158Leu	57	1229	4.6									
	PTEN	p.Ile4Thr	45	1002	4.5									
	MAP2K1	p.Ser72Gly	24	629	3.8									

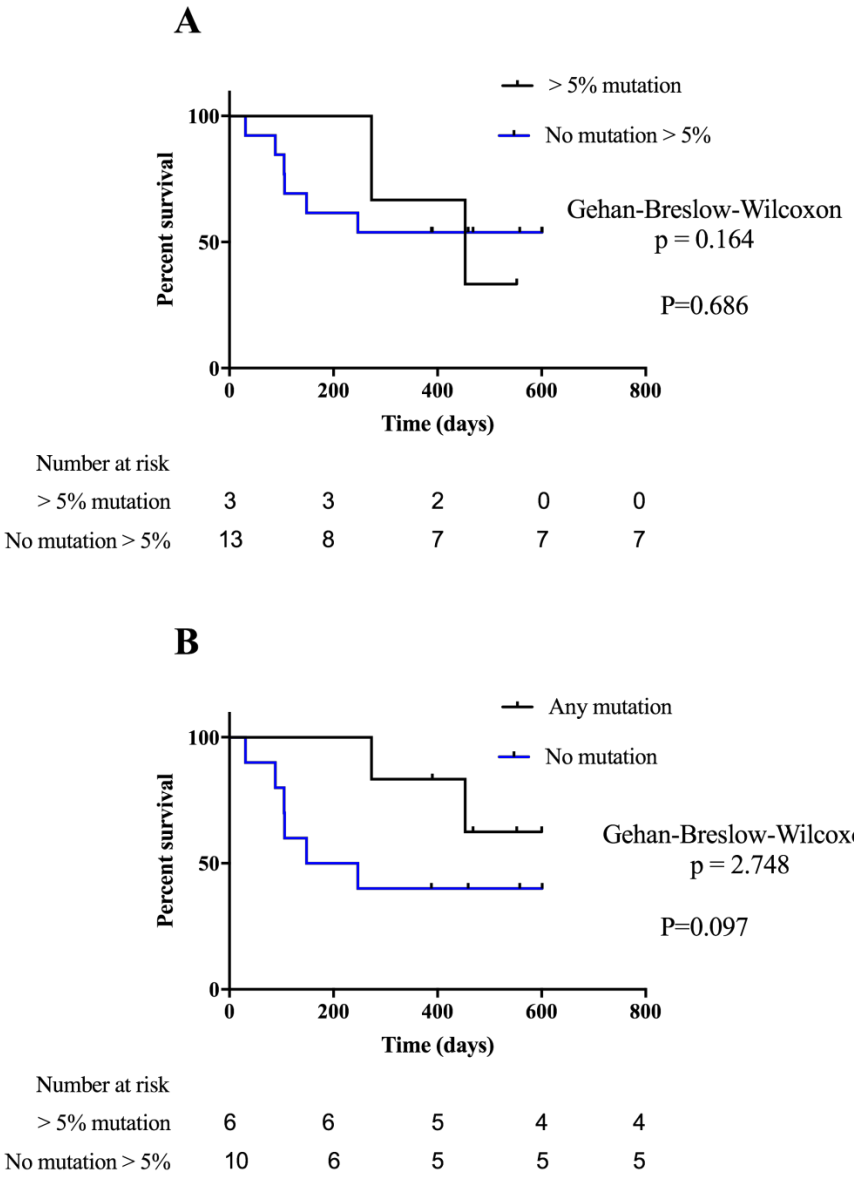
Pt No	Mutation		Blood			Primary			Lymph node			Metastatic		
	Gene	Protein	Variant	total	%	Variant	total	%	Variant	total	%	Variant	total	%
10	TP53	p.Arg333Cys	60	1009	5.9	0	967	0						
		p.Met237Val	57	1160	4.9	0	1087	0	0	928	0			
		p.Arg306Gly	39	1028	3.8	0	976	0	0	971	0			
		p.Val197Met	19	658	2.9	0	317	0						
	SMAD4	p.Gly510Glu	61	304	20									
		p.Cys401Arg	30	705	4.3									
	DDR2	p.Val100Met	92	646	14									
	STK11	p.Asp348Gly	32	512	6.3									
	ALK	p.Gly1202Arg	78	1465	5.3									
	NRAS	p.Asn116Asp	7	152	4.6									
	ERBB4	p.Ser303Tyr	46	1320	3.5									
	PIK3CA	p.His701Arg	44	1317	3.3	0	2907		0	3996				
	FGFR3	p.Gly805Ser	9	272	3.3									
13	STK11	p.Ala200Val	13	316	4.1	0	68		0	202				
16	AKT1	p.Pro51Ser	35	893	3.9				0	264				

Table 6-12 continued from previous page

6.2.6 Survival analysis

Identification of mutations over the 5% threshold or of ANY frequency was not associated with a difference in overall survival Error! Reference source not found..

Figure 6-6 Kaplan-Meier analysis for overall survival of OncoQuik® cohort. A) Comparison of patients with mutations identified over 5% threshold versus those patients with no mutations over 5%. B) Comparision of patients with mutations identified at ANY threshold.



6.2.7 Patient characteristics of cancer and control patients in Rosette Sep™ cohort

The characteristics of cancer and control patients in the Rosette Sep™ cohort are shown in Table 6-13. As patients in the RosetteSep™ cohort were recruited on an individual basis by the investigator and not from the generic PBRU Biobank, complete demographic information is not available.

Table 6-13. Characteristics of cancer and control patients in the Rosette Sep™ cohort

*Stage reported according to histology if resected or according to radiology if not resected.

There were significantly more patients with variants of greater than 5% frequency in the cancer group compared to the control group (P=0.005) Table 6-14. There were also significantly more patients with variants at any frequency in the cancer compared to the control group (P=0.0215)
Table 6-15 and Figure 6-7.

Table 6-14. Compariosn of mutations >5% in cancer and control patients in the Rosette Sep™ cohort. P=0.005 using Fiseher's exact

	>5% mutation	No mutation >5 %	Total
Cancer	8	7	15
Control	0	9	9
Total	8	16	24

Table 6-15. Comparison of mutations at ANY frequency in cancer and control patients in the Rosette Sep™ cohort. P=0.0215 using Fiseher's exact

	Any mutation	No mutation	Total
Cancer	13	2	15
Control	3	6	9
Total	16	8	24

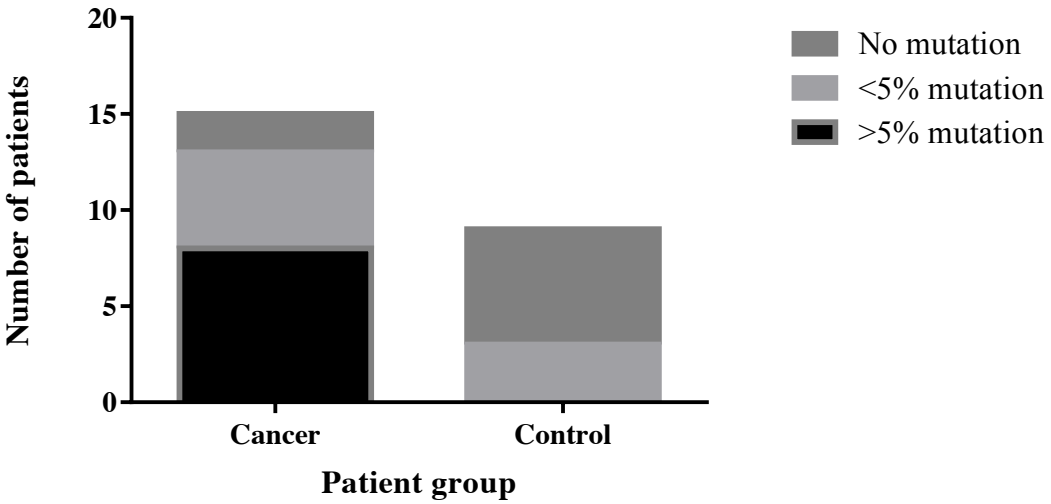


Figure 6-7. Comparison of mutation frequency identified between cancer and control groups in the Rosette Sep™ cohort

One control patient was excluded from analysis due to the identification of a previously unknown germline mutation (p.Arg283Cys) in the TP53 gene, Li-Fraumeni Syndrome. This patient however did not have any mutations identified over the 5% threshold with the exception of the germline TP53 described. A mutation with a frequency greater than 5% identified in the control group was excluded on the basis that it was an INDEL (EGFR: c.1431_1432insA) in a high homopolymer region of the epidermal growth factor receptor (EGFR) gene with 6 consecutive ‘A’s. The ion torrent platform is notoriously susceptible to miscalling in this setting and therefore the mutation is assumed to be a false positive Figure 6-8.

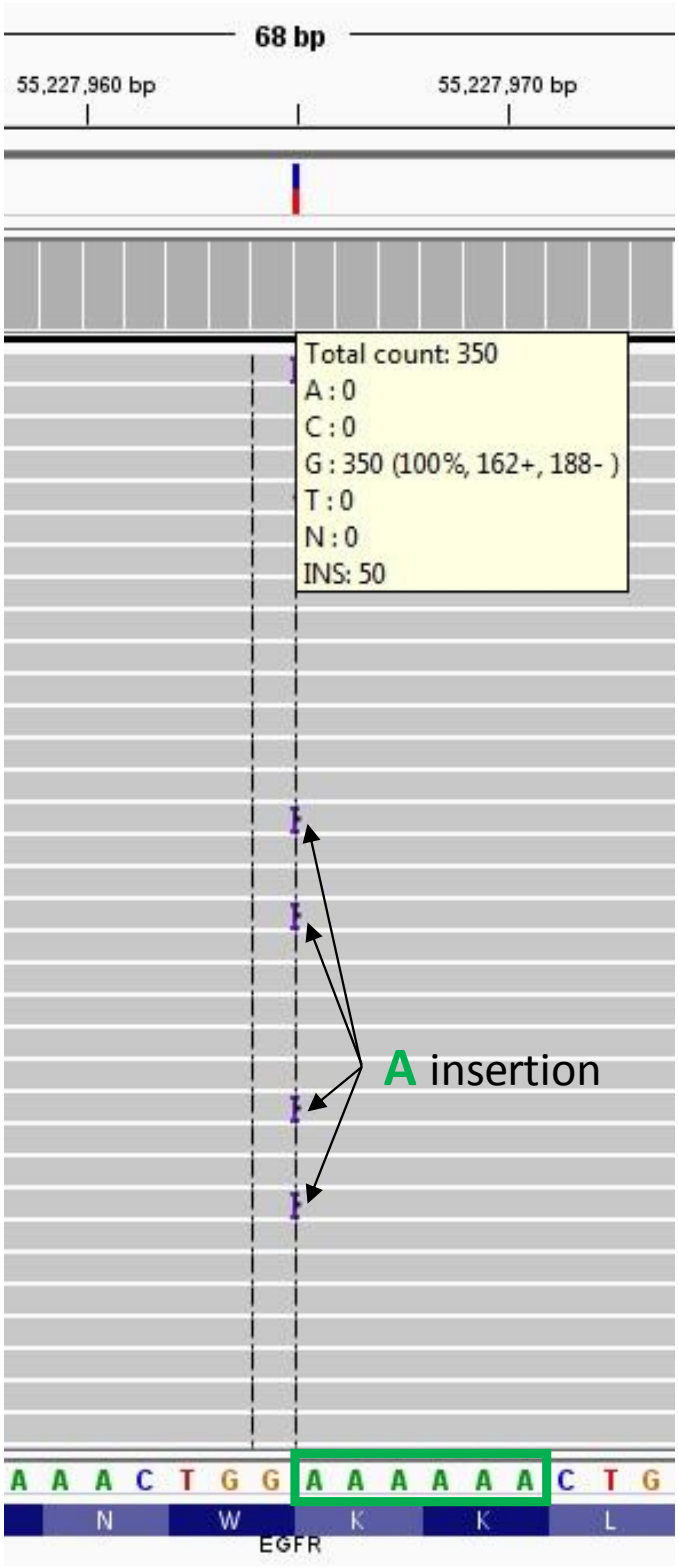


Figure 6-8. Mutation in the EGFR gene identified in a control patient and subsequently excluded from analysis. The insertion of an 'A' base pair after a run of 6 consecutive 'A's is most likely due to base miss-calling by the ion torrent platform.

6.2.8 Analysis of the mutations identified by Rosette Sep™ enrichment method.

RosetteSep™ CTC Enrichment Cocktail Containing Anti-CD36 (RS-36) performed the best overall identifying 6 mutations with a frequency >5% in 5 patients Figure 6-9. Comparison of mutations >5% identified according to Rosette SepTM enrichment method

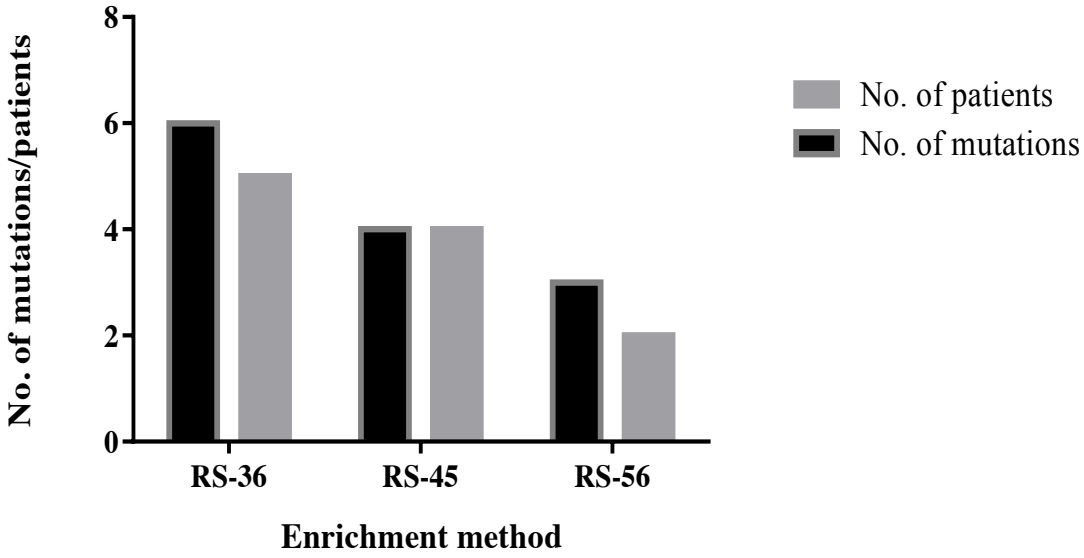


Figure 6-9. Comparison of mutations >5% identified according to Rosette Sep™ enrichment method

6.2.9 Comparison of mutations from Rosette Sep™ enriched blood with primary tumour.

Of the seven patients who underwent resection in the Rosette Sep™ cancer cohort only one (patient 6) had FFPE tumour material available for sequencing. This sample was a moderate and focally poorly differentiated PDAC and staged at pT3,N1,M0,R1. *KRAS* and T*P53* mutations were identified; p.Gly12Cys and p.Gly199Val respectively. These mutations were not identified in any of the RS enriched blood samples for that patient. Similarly, none of the 8 variants identified in the RS enriched blood sample were found in the primary tumour.

6.2.10 Tracking mutational profile across serial blood samples.

Serial blood samples were acquired from 10 of the 15 cancer patients. Three samples were taken in total from each patient during the course of their treatment. All samples were separated by administration of chemotherapy representing a potential clonal sweep of CTCs. The treatment timelines for the 10 patients indicating the timing of the sample collection in relation to their treatment is shown in Figure 6-10. The variants identified and their presence or absence across the three separate samples is shown in Figure 6-11. Of the 47 variants identified, 37 were identified exclusively in the first sample (S1), 8 exclusively in the second sample (S2), 1 exclusively in the third sample (S3) and 1 was identified in all three samples. With the exception of this trans-sample mutation, none of the mutations identified were present in both duplicates for S1 samples or present in the other two corresponding RS enrichments.

In patient 3 the *TP53* mutation p.Arg267Trp was identified in all three serial samples (S1,S2 and S3), and, in the second sample (S2), in all three RS enrichment samples (albeit at low coverage in RS-56 and RS-36) (Figure 6-12). The principle identification of this variant was in the patient's 1st sample (S1), RS-56 enrichment in the second duplicate (Figure 6-12A). Here, a very high Phred quality score (Q) of 250.2 infers the probability that this variant was incorrectly called due to sequencing error is 1 in $1x10^{25}$. The Phred quality scores for this variant in the subsequent samples (Figure 6-12B and C) were 111.1, 10.6, 11.96, 7.405 and 103.6. The SIFT score for this variant is 0.0 suggesting the amino acid change has a 'deleterious' effect on *TP53* function. The variant has been confirmed as a somatic mutation in the COSMIC database with reference COSM11183 and the IARC TP53 database which has compiled 33 reported cases in the literature of this mutation in association with cancer.

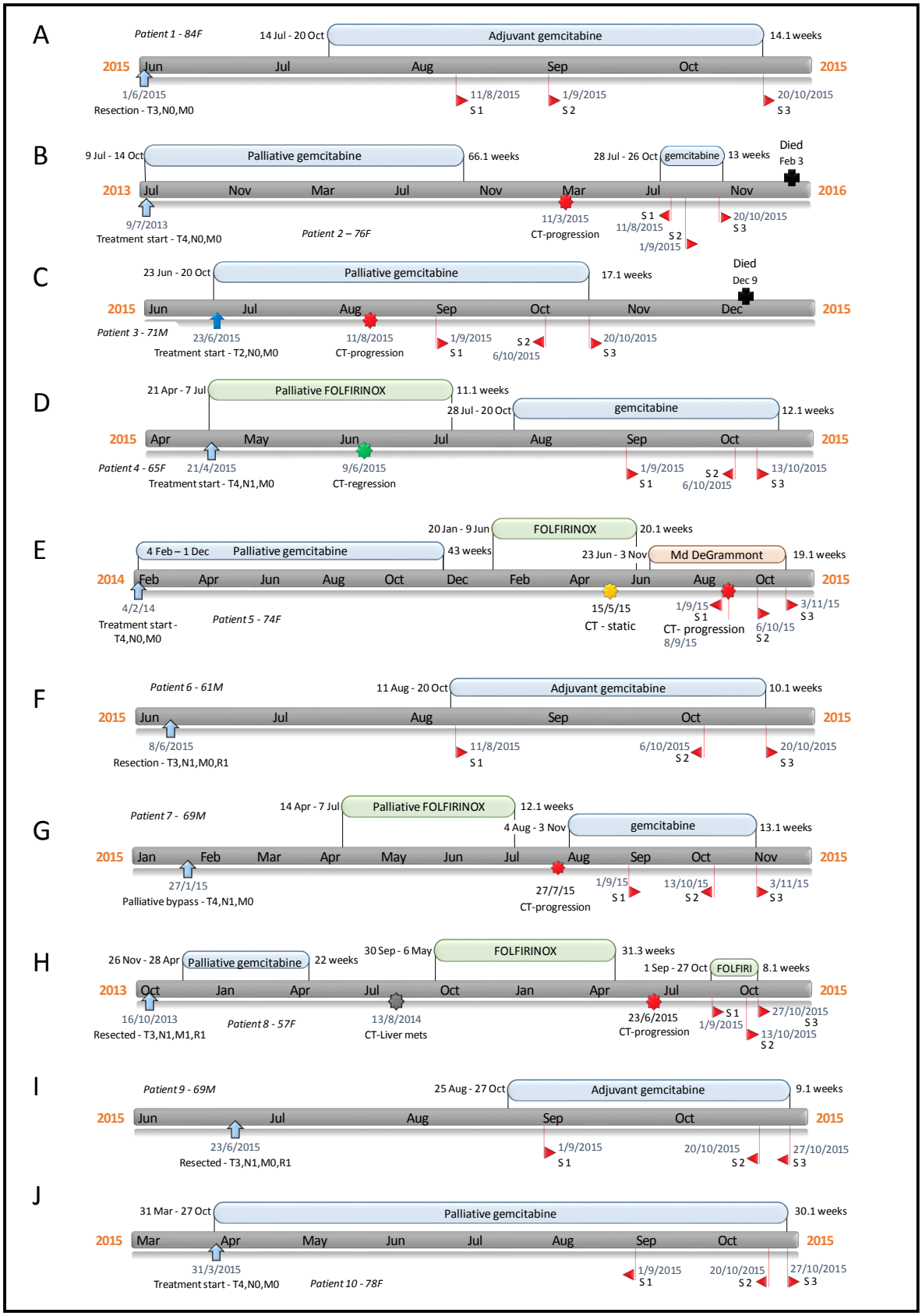


Figure 6-10. Treatment and blood sampling timeline of 10 cancer patients in Rosette Sep™ cohort undergoing serial sampling. A) patient 1, B) patient 2 etc. through to J) patient 10. S1, S2 and S3 represent the first, second and third

sampling of blood for Rosette Sep enrichment respectively. Gemcitabine chemotherapy given in 28-day cycle with treatment on days 1, 8 and 15. FOLFIRINOX (Oxaliplatin, Irinotecan and Fluorouracil) and FOLFIRI (Irinotecan and Fluorouracil) given in 14 day cycle with treatment on day 1 and 5FU pump for 46hrs. Md Grammont - Modified DeGrammont (Fluorouracil) given in 14 day cycle with treatment on day 1 and 5FU pump for 46hrs.

Pt No.	Gene	Mutation	RS™	S1	S2	S3	Pt No.	Gene	Mutation	S1	S2	S3
1	*TP53*	p.Ser303Asn		3.6			4	*FGFR3*	p.Pro696Ser	4.4		
	NRAS	p.Pro140Leu		4.4				*FGFR2*	p.Asn550fs		5.4	
		p.Gly138Glu		3.8				*DDR2*	p.Arg473His	6.7		
		p.His131Tyr		4.6			5	*TP53*	p.Arg248Trp	3.9		
	FGFR1	p.Gly301Asp		3.4					p.Lys132Arg	2.3		
2	*TP53*	p.Pro152Leu		2.8			6	*TP53*	p.Arg248Gln	3.4		
		p.Pro300Ser		2.5					p.Glu286Lys	2.9		
		p.Pro82Leu			3			*FGFR3*	p.Gly712Asp	4.5		
		p.Arg110His			7.1			*NRAS*	p.Glu132Gly	4		
	ERBB4	p.Arg232Ter		3.4				*FBXW7*	p.Met268Ile	4.1		
	PIKC3A	p.Met1043Val		2.31				*ERBB4*	p.Arg612Trp	3.7		
	PTEN	p.Lys66Glu			2.7			*BRAF*	p.Gly593Ser	8.1		
	SMAD4	p.Ala451Val		4.11			7	*TP53*	p.Pro219Leu	4.2		
		p.Arg189His			4.4			*STK11*	p.His27Tyr	3.9		
	DDR2	p.Gly108Arg		7.6			8	*FBXW7*	p.Glu452Lys	6.2		
	ALK	p.Arg1275Gln		4.7				*ERBB4*	p.Gly240Glu	3.73		
	MET	p.Asn375Ser			4.8		9	*TP53*	p.Arg181Cys	2.9		
	FGFR3	p.His251Tyr			4.4		10	*TP53*	p.Arg181Cys	2.80		
	STK11	p.Gly346Ser			6.4			*FBXW7*	p.Arg505Leu	2.7		
3	*TP53*	p.Arg267Trp	56	6.6	6.3			*ERBB4*	p.Asp922Asn	4.1		
			45	4.8	2.2	3.5		*SMAD4*	p.Trp101Arg			4.3
			36		6.4			*BRAF*	p.Asp587Asn	3.7		
		p.Asp208Asn		2.4				*FGFR3*	p.Ala265Val	3.3		
	ERBB4	p.Leu139Ser		5.4			Key:		Mutation identified %			
		p.Glu141Lys		3.8					No mutation identified			
	SMAD4	p.Ala258Val		3.7					Failed sequencing			

Figure 6-11. Tracking of mutations identified in Rosette Sep™ enriched blood samples taken at three separate time points during the course of patients' chemotherapy treatment for PDAC.

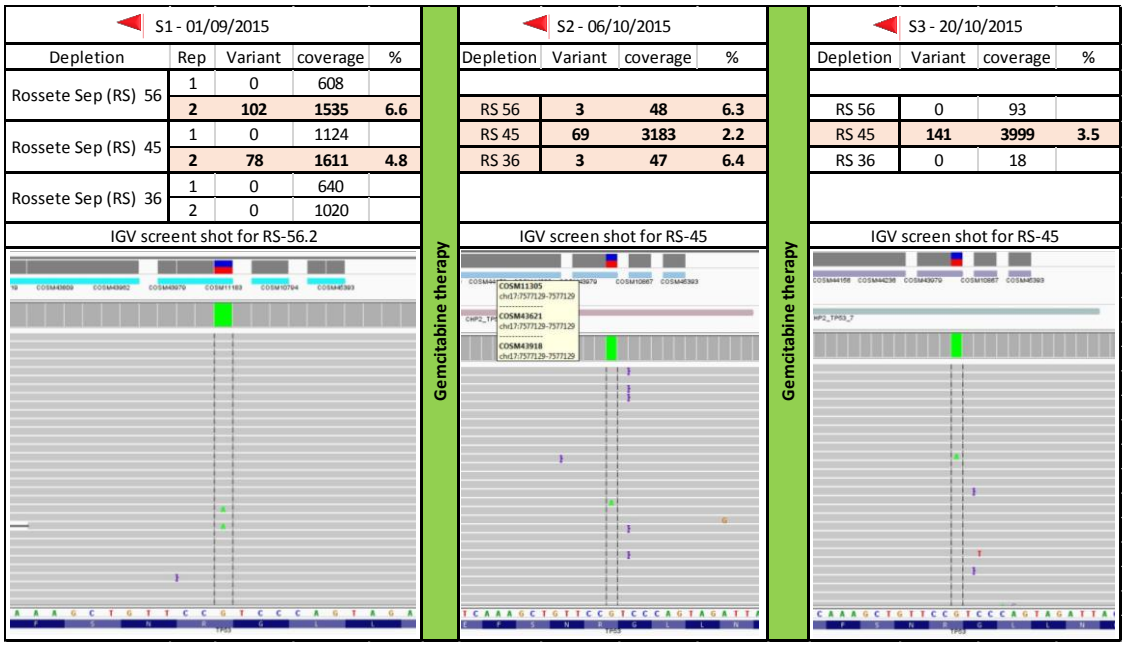


Figure 6-12. The *TP53* mutation p.Arg267Trp identified in all three serial samples from patient 3.

6.3 Discussion

In this chapter the two most promising CTC enrichment technologies OncoQuik® and Rosette Sep™ were taken forward and applied to two separate cohorts of patients, those undergoing pancreatic surgery for PDAC and those receiving chemotherapy for PDAC respectively.

The first aspect of this work involved switching from 'in-house' primers, covering just exons, 5 to 8 of the *TP53* gene, to a panel covering the hotspots of 22 cancer related genes. PDAC is well represented in this panel with all the mountain mutations with the exception of *CDKN2A* covered *(TP53, KRAS* and *SMAD4*) and a large number of lower frequency mutations, comprising the long genomic tail of PDAC as shown in The Oncomine™ panel contains multiplex primers for the hotspots of 22 frequency mutated genes. A list of the genes and their reported frequency in PDAC is listed in Table 2-5**Error! Reference source not found.**. A number of the genes have actionable phenotypes vastly increasing the scope for personalised therapy compared to sequencing *TP53* alone. The PDAC cell line spiking experiments were an important step to confirm that an adequate depth >1,000 reads could be achieved within the new sequencing parameters. It also confirmed that common mutations in both *TP53* and *KRAS* could be accurately and consistently identified.

Differences in tumour stages, treatments and sample processing methods make direct comparison between the OncoQuik® and RosetteSep™ cohort difficult, though this was not the aim of the thesis. The RosetteSep™ Cohort identified significantly more variants both above and below the 5% frequency threshold when compared to age matched controls. Comparison with primary tumour was only possible in one patient in this cohort, and, as with the OncoQuik® cohort no variants were found in both. Also concerning is the fact that the nearly ubiquitous KRAS mutations (e.g. p.G12D or p.G12C) were not found in any of the samples which would be expected if their origin was PDAC.

In patient 2 an additional mutation was superimposed in the lymph node (*TP53* p.Arg196Ter) which was not present in the primary tumour. This could be an example of branched evolution, with p.Arg196Ter being 'private' (i.e. late) mutation(Gerlinger et al., 2012)

The profiles of the 10 serially sampled patients in the Rosette Sep™ cohort shows that variants emerge and drop out at all time points. This is consistent with theories that clonal diversity is driven by branched tumour evolution, responding to selective pressure from the local microenvironment including chemotherapy(Burrell et al., 2013; Greaves & Maley, 2012b). Though the first S1 stage identified most, and the S3 stage least variants. Variant changes were seen within the same gene, for example different variants were detected in *TP53* and *SMAD4* between the first and second samples in patient 2 Figure 6-11. Similarly, significant genetic disparity, so called 'micro heterogeneity' has been demonstrated between positively (EpCam) selected CTCs in cancer patients(Polzer et al., 2014). Whilst the origin of these variants cannot be proven due to the nature of the negative selection enrichment it is feasible that these changes reflect changes of the dominant clone of CTC following clonal sweeps with chemotherapy treatment. The clinical application of this is potentially large with personal selection of chemotherapy being possible based on the molecular profile of the dominant clone at the time of treatment and not that of the primary tumour or even initial CTC profile.

Additionally, it was shown that a somatic, deleterious mutation widely implicated in carcinogenesis(*TP53*, p.Arg267Trp) (AlHarbi et al., 2018) could be identified in the peripheral blood without a priori knowledge, and subsequently tracked across three serial samples during the course of treatment. This mutation was detected in all three RS enrichment methods and satisfied all the requirements for a disease causing mutation. That it was not detected in both replicates of the first S1 sample also adds weight the utility of the limiting dilution method used. A germline variant would be expected to be present in every sample. A sequencing artefact would not be reproduced in multiple samples of only one patient in the low homopolymer regions. This phenomena of tracked mutation throughout a patient's samples was only seen in this one patient. One may speculate that the presence of disease progression whilst on chemotherapy, and death due to cancer metastasis less than 50 days after the third sample (S3) (Figure 6-10C) observed in this patient is a reflection of a higher burden of CTCs thus permitting detection and tracking. If this is the case, and the issue is remains one of contaminating wild type DNA, then perhaps reverting to 10 genomes per/µL with 10 replicates (as opposed to two duplicates and 20 genomes per/µL) may overcome this

issue. The importance of this finding is that it is the first step towards the ultimate goal of personalised therapy based on CTCs rather than primary tumour.

Other cancer types have successfully used personalised therapy based on the molecular analysis of the primary tumour; it has been the mainstay of breast cancer treatment since the 1980s(Ross et al., 2009) and lung(Kim et al., 2011; Pao & Chmielecki, 2010), colorectal(Kocarnik, Shiovitz, & Phipps, 2015) and prostate cancer(Den et al., 2015) appear set to benefit in the near future. Unfortunately, efforts to date of personalised therapy in PDAC have been unsuccessful. The Individualised Molecular Pancreatic Cancer Therapy (IMPaCT) study randomised patients between standard chemotherapy or personalised chemotherapy based on 4 sub-groups of actionable mutations uncovered by sequencing of the primary tumour. A pilot study failed to recruit a single patient, and serves to emphasise the difficulties: that poor quality, inaccessible, untimely, heterogeneous bio-specimens would make molecular characterisation of the primary tumour to guide therapy impractical in an adequately powered trial(Chantrill et al., 2015). The National Cancer Institute's Molecular Analysis for Therapy Choice (NCI-MATCH) studies include pancreatic cancer as a priority area and have shown the feasibility of recruitment of pancreatic cancer patient for clinical trials of precision medicine, albeit that to date no clinical benefit has been shown (Coyne, Takebe, & Chen, 2017). There are numerous barriers to personalised therapy in PDAC related to genetic heterogeneity, clonal evolution, and biospecimen acquisition and composition. Unfortunately, because of these barriers the use of personalised therapy in pancreatic cancer remains at the proof of concept stage. The ability to extract relevant genomic material from the CTCs in patients with pancreatic cancer is key to reaching the “holy grail” of personalised therapy.

7 DISCUSSION

The first aim of the thesis was to develop and optimise a technology to enrich and molecularly characterise CTCs in PDAC. To this end we began with examining the molecular changes in pancreatic juice. The EUROPAC group, with which I was affiliated had already established a research programme of pancreatic juice analysis for risk stratification in patients with a genetic predisposition for PDAC(Grocock et al., 2007). This was an opportunity to develop downstream molecular analysis which could later be applied to CTCs as well as optimise pancreatic juice as a source of biospecimen. We established that whole juice was superior to supernatant in discriminating PDAC from control samples and no evidence was found that pellet was inferior to whole juice. The sensitivity and specificity of *KRAS* status for PDAC in the pellet of pancreatic juice was 83.3% and 64.3% respectively.

We then turned to developing an in-house method of CTC enrichment using negative depletion combined with low frequency variant detection using NGS and limiting dilution. In various small-scale spiking experiments, a single CD45 depletion was found to be the optimal strategy for CD45 depletion and ultimately CTC enrichment. This was then taken forward for further assessment of four clinical samples. Developing the technology using clinical volumes (7mLs) of blood required an exponential scale-up of reagents. For example, a significantly larger volume of reagents (antibodies and ferrous beads etc) were required for each depletion compared to earlier spiking experiments. Each patient also had typically three additional experimental controls with each of these undergoing multiple repeats (up to 10), which were then sequenced using in-house *TP53* primers and NGS technology. The in-house primers did not use the multiplex PCR methods of the Oncomine™ panel used later in the thesis and was therefore an extremely labour-intensive process. The implication of this arrangement was that it was not feasible to increase the number of patients beyond four given the limited resources of time and money available. Whilst the in-house CD45 depletion and *TP53* primers showed promise and appeared to significantly enrich CTCs in half of the patients, there were some inconsistencies. We therefore decided not to take this technology into the next phase of development and instead looked towards established CTC technologies to develop further.

Three commercial CTC enrichment technologies were selected for assessment to cover a variety of enrichment methods; CellSieve™ using size, OncoQuik® using buoyancy and RosetteSep™ using negative selection. In several clinical samples using appropriate

controls, repeats and head to head comparisons OncoQuik® and RosetteSep™ out performed CellSieve™. The majority (8/9) of *TP53* mutations were identified in the non-enriched portion <10μm using CellSieve™ suggesting that a smaller phenotype of CTCs capable of passing through the filter may predominate in these samples consistent with recent reports of EMT(Brabletz et al., 2018)(Alix-Panabières, Mader, & Pantel, 2017). RosetteSep™ and OncoQuik® showed promise with consistency both within and across enrichment methods. For example, in patient 11 the *TP53* non-synonymous deleterious mutation p.L265Q was identified using OncoQuik® (all 5 repeats at a mutation frequency of 12%), RosetteSep™-56 (3 out of 5 repeats at a mutation frequency of 6.8%) and RoesetteSep™-36 (1 out of 5 repeats at a mutation frequency of 9.9%) and subsequently confirmed using mutation specific PCR. These results satisfy the first aim of the thesis to develop and optimise a technology to enrich and molecularly characterise CTCs in PDAC.

The second aim of the thesis was to use the technology developed to track the mutational profile of CTCs through a patient's treatment. In the final chapter this was attempted using two different cohorts of patients (OncoQuik® and RosetteSep™) with PDAC and appropriate controls. The OncoQuik® enrichment was performed pre-operatively on the day of pancreatic surgery on 29 patients. 16 patients underwent surgery for PDAC in the cancer group and 13 patients underwent surgery for benign conditions in the benign group.

An attempt was made track mutations in the OncoQuik® cohort from pre-operative sampling through to post-operative sampling. Following up patients for serial sampling of blood proved a significant logistical challenge. Some patients had follow-up at their referring hospital which I did not have access to, and on many occasions, technical support to process the samples fresh on the day of clinic follow up was not available. Consequently, in the OncoQuik® cohort only four patients underwent sequencing of enriched CTCs in the post-operative follow up period, and unfortunately, three of these patients were excluded from analysis as they had non-PDAC pathology. Clearly it is not possible to draw conclusions from such limited data. Whilst this follow up sequencing data would have been interesting, lack of it does not detract from the aims of the thesis which was achieved in the RosetteSep™ cohort.

Several questions were raised in this chapter about the ability of OncoQuik® to enrich whole blood for CTCs. Only a trend towards significance once the 5% variant frequency threshold had been removed was found and there was not enough data to show significance with the 5% threshold. There was also no correlation between the variants identified in the blood and that of the solid tumour material (albeit that the sequence data on the tissue at the corresponding sites was in most cases poor). The mutation status was not shown to have a bearing on overall survival with a trend for better survival with samples where mutation positive CTCs were isolated. There are several potential explanations for this. The OncoQuik® contained patients with the earliest stages of various cancers when levels of CTCs are expected to be at their lowest, perhaps below the threshold for detection at this stage. The freezing and storing process permitted batch processing of these relatively rare samples which took many months to obtain. However, this process may have reduced the viability of CTCs or altered their buoyancy such that the enrichment process became ineffective. Perhaps the buoyancy of CTCs in early cancers is different to that found in more advanced cancers and hence evaded isolation. One drawback of this method of negative selection of CTCs is that it is not possible to confirm the presence of CTCs in the sample before processing. It is possible therefore that the variants identified were from other sources of DNA such as circulating free DNA (cfDNA) found in plasma: although by pelting the cells before processing most of the plasma should have been discarded, it is inevitable that a small volume will get carried over with the pellet. A number of studies have described the use of cfDNA to screen for cancer(Bianchi et al., 2015; Heitzer, Ulz, & Geigl, 2015; Szpechcinski et al., 2015), recurrence(Bratman, Newman, Alizadeh, & Diehn, 2015; Olsson et al., 2015; Stremitzer et al., 2015)and response. Plasma contains approximately 1µg/ml of free DNA(Bagul, Pushpakom, Boylan, Newman, & Siriwardena, 2006), most comes from leukocytes and endothelial cells, but in cancer patients the levels can rise by as much as 10 fold; even more during chemo and radio therapy(Holdenrieder et al., 2001). Some of this increase may be due to release of DNA from lysed apoptotic or necrotic tumour cells but the largest proportion results from active secretion from macrophages; work in mouse models suggests that this cancer induced increase includes nucleosomes that have not come from cancer cells(Pisetsky, 2004).

In the second cohort of patients RosetteSep™ was used to enrich CTCs and included 15 patients with PDAC undergoing chemotherapy (10 of whom were sampled serially on three separate occasions between cycles of chemotherapy) and 10 controls patients

without cancer. This group represented a significantly easier group to serially sample as their timings of chemotherapy treatments offered an ideal opportunity to take repeat samples. The RosetteSep^{TM} samples were also able to be processed fresh and so I was not limited by the availability of technical support staff.

In the final chapter I demonstrated that it was possible to track a non-synonymous, deleterious *TP53* mutation (p.R267W) identified in an enriched blood sample across three serial time-points in the patient's treatment; thus satisfying the second aim of the thesis: to use the CTC and NGS technology developed to track the mutational profile of CTCs through a patient's treatment.

I recommend that future research be directed towards determining the prognostic significance mutations in CTCs persisting during treatment with a larger and more homogeneous cohort of PDAC patients. Having established an optimal enrichment method and suitable downstream analysis, resources in these studies could be more focused than that used here in the development phase. This would permit each sample to have a greater number of serial repeats, permitting a more dilute concentration of sample and potentially increasing the sensitivity of mutation detection further. If future trials demonstrated the prognostic significance of the persistence of CTCs during chemotherapy treatment then further clinical trials could examine whether switching chemotherapy agents in this scenario (a potential surrogate marker of chemo resistance) improved overall survival.

Personalised therapy is the natural progression from targeted therapy, whereby the targeted therapy is selected for a patient based on actionable mutations found in the patient's tumour(Tran et al., 2016). Ultimately the goal would be to use the molecular profile of CTCs in a PDAC patient upfront to select the most appropriate chemotherapy agent, and then switch agents when chemoresistance is identified though there are many barriers including CTC enrichment technology, genetic heterogeneity, clonal evolution and biospecimen type which need to be overcome first.

8 REFERENCES

Adams, D. L., Martin, S. S., Alpaugh, R. K., Charpentier, M., Tsai, S., Bergan, R. C., … Cristofanilli, M. (2014). *Circulating giant macrophages as a potential biomarker of solid tumors*. https://doi.org/10.1073/pnas.1320198111

Adams, D., Zhu, P., Makarova, O. V, Martin, S., Charpentier, M., Chumsri, S., … Tang, C. (2015). The systematic study of circulating tumor cell isolation using lithographic microfilters. *Rsc Advances*, *9*, 4334–4342. https://doi.org/10.1039/c3ra46839a

AlHarbi, M., Mubarak, N., AlMubarak, L., Aljelaify, R., AlSaeed, M., Almutairi, A., … Abedalthagafi, M. (2018). Rare TP53 variant associated with Li-Fraumeni syndrome exhibits variable penetrance in a Saudi family. *Npj Genomic Medicine*, *3*(1). https://doi.org/10.1038/s41525-018-0074-3

Alix-Panabières, C., Mader, S., & Pantel, K. (2017). Epithelial-mesenchymal plasticity in circulating tumor cells. *Journal of Molecular Medicine*, *95*(2), 133–142. https://doi.org/10.1007/s00109-016-1500-6

Allard, W. J., Matera, J., Miller, M. C., Repollet, M., Connelly, M. C., Rao, C., … Terstappen, L. W. M. M. (2005). *Tumor Cells Circulate in the Peripheral Blood of All Major Carcinomas but not in Healthy Subjects or Patients With Nonmalignant Diseases Tumor Cells Circulate in the Peripheral Blood of All Major Carcinomas but not in Healthy Subjects or Patients With Non*. 6897–6904.

Altin, J. G., & Sloan, E. K. (1997). *The role of CD45 and CD45-associated molecules in T cell activation*. 430–445.

Andea, A., Sarkar, F., & Adsay, V. N. (n.d.). *Clinicopathological Correlates of Pancreatic Intraepithelial Neoplasia : A Comparative Analysis of 82 Cases With and 152 Cases Without Pancreatic Ductal Adenocarcinoma*. (3). https://doi.org/10.1097/01.MP.0000087422.24733.62

Andreopoulou, E., Yang, L. Y., Rangel, K. M., Reuben, J. M., Hsu, L., Krishnamurthy, S., … Cristofanilli, M. (2012). Comparison of assay methods for detection of circulating tumor cells in metastatic breast cancer: AdnaGen AdnaTest BreastCancer Select/Detect versus Veridex CellSearch system. *International Journal of Cancer. Journal International Du Cancer*, *130*(7), 1590–1597. https://doi.org/10.1002/ijc.26111 [doi]

Armstrong, A. J., Marengo, M. S., Oltean, S., Kemeny, G., Bitting, R. L., Turnbull, J. D., … Garcia-Blanco, M. a. (2011). Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. *Molecular Cancer Research : MCR*, *9*, 997–1007. https://doi.org/10.1158/1541-7786.MCR-10-0490

Arumugam, T., Ramachandran, V., Fournier, K. F., Wang, H., Marquis, L., Abbruzzese, J. L., … Choi, W. (2009). Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Research*, *69*(14), 5820–5828. https://doi.org/10.1158/0008-5472.CAN-08-2819

Asano, T., Yao, Y., Zhu, J., Li, D., Abbruzzese, J. L., & Reddy, S. a. (2005). The rapamycin analog CCI-779 is a potent inhibitor of pancreatic cancer cell proliferation. *Biochemical and Biophysical Research Communications*, *331*(1), 295–302. https://doi.org/10.1016/j.bbrc.2005.03.166

Auerbach, C., Moutschen-Dahmen, M., & Moutschen, J. (1977). Genetic and cytogenetical effects of formaldehyde and related compounds. *Mutation Research*, *39*(3–4), 317–361. https://doi.org/10.1016/0165-1110(77)90011-2

Bagul, A., Pushpakom, S., Boylan, J., Newman, W., & Siriwardena, A. K. (2006). Quantitative analysis of plasma DNA in severe acute pancreatitis. *JOP : Journal of the Pancreas*, *7*(6), 602–607.

Balic, M., Dandachi, N., Hofmann, G., Samonigg, H., Loibner, H., Obwaller, A., … Bauernhofer, T. (2005). Comparison of two methods for enumerating circulating tumor cells in carcinoma patients. *Cytometry.Part B, Clinical Cytometry*, *68*(1), 25–30. https://doi.org/10.1002/cyto.b.20065 [doi]

Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., … Wooster, R. (2004). The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer*, *91*(2), 355–358. https://doi.org/10.1038/sj.bjc.6601894

Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K. W., Velculescu, V. E., … Nowak, M. A. (2007). Genetic progression and the waiting time to cancer. *PLoS Computational Biology*, *3*(11), 2239–2246. https://doi.org/10.1371/journal.pcbi.0030225

Bianchi, D. W., Chudova, D., Sehnert, A. J., Bhatt, S., Murray, K., Prosen, T. L., … Halks-Miller, M. (2015). Noninvasive Prenatal Testing and Incidental Detection of Occult Maternal Malignancies. *Jama*, *314*(2), 162. https://doi.org/10.1001/jama.2015.7120

Biankin, A. V., Waddell, N., Kassahn, K. S., Gingras, M.-C., Muthuswamy, L. B.,

Johns, A. L., … Grimmond, S. M. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature*, *491*(7424), 399–405. https://doi.org/10.1038/nature11547

Bidard, F. C., Huguet, F., Louvet, C., Mineur, L., Bouche, O., Chibaudel, B., … Hammel, P. (2013). Circulating tumor cells in locally advanced pancreatic adenocarcinoma: the ancillary CirCe 07 study to the LAP 07 trial. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO*, *24*(8), 2057–2061. https://doi.org/10.1093/annonc/mdt176 [doi]

Bondar, V. M., Sweeney-Gotsch, B., Andreeff, M., Mills, G. B., & McConkey, D. J. (2002). Inhibition of the phosphatidylinositol 3'-kinase-AKT pathway induces apoptosis in pancreatic carcinoma cells in vitro and in vivo. *Mol Cancer Ther*, *1*(12), 989–997. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12481421

Brabletz, T., Kalluri, R., Nieto, M. A., & Weinberg, R. A. (2018). EMT in cancer metastasis. *Nature Reviews Cancer*, *18*, 129–135.

Bratman, S. V, Newman, A. M., Alizadeh, A. a, & Diehn, M. (2015). Potential clinical utility of ultrasensitive circulating tumor DNA detection with CAPP-Seq. *Expert Review of Molecular Diagnostics*, *15*(6), 1–5. https://doi.org/10.1586/14737159.2015.1019476

Bruns, C. J. (2004). Rapamycin-Induced Endothelial Cell Death and Tumor Vessel Thrombosis Potentiate Cytotoxic Therapy against Pancreatic Cancer. *Clinical Cancer Research*, *10*(6), 2109–2119. https://doi.org/10.1158/1078-0432.CCR-03-0502

Burrell, R. a, McGranahan, N., Bartek, J., & Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*, *501*(7467), 338–345. https://doi.org/10.1038/nature12625

Campbell, P. J., Yachida, S., Mudie, L. J., Stephens, P. J., Pleasance, E. D., Stebbings, L. a, … Futreal, P. A. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature*, *467*(7319), 1109–1113. https://doi.org/10.1038/nature09460

Canto, M. I., Harinck, F., Hruban, R. H., Offerhaus, G. J., Poley, J.-W., Kamel, I., … Bruno, M. (2013). International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. *Gut*, *62*(3), 339–347. https://doi.org/10.1136/gutjnl-2012-

303108

Castellanos, J. a., Merchant, N. B., & Nagathihalli, N. S. (2013). Emerging targets in pancreatic cancer: Epithelial-mesenchymal transition and cancer stem cells. *OncoTargets and Therapy*, *6*, 1261–1267. https://doi.org/10.2147/OTT.S34670

Chantrill, L. a, Nagrial, A. M., Watson, C., Johns, A. L., Martyn-Smith, M., Simpson, S., … Biankin, A. V. (2015). Precision Medicine for Advanced Pancreas Cancer: The Individualized Molecular Pancreatic Cancer Therapy (IMPaCT) Trial. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *21*(9), 2029–2037. https://doi.org/10.1158/1078-0432.CCR-15-0426

Cheng, C.-L., Sherman, S., Watkins, J. L., Barnett, J., Freeman, M., Geenen, J., … Lehman, G. a. (2006). Risk factors for post-ERCP pancreatitis: a prospective multicenter study. *The American Journal of Gastroenterology*, *101*(1), 139–147. https://doi.org/10.1111/j.1572-0241.2006.00380.x

Cohen, S. J., Punt, C. J., Iannotti, N., Saidman, B. H., Sabbath, K. D., Gabrail, N. Y., … Meropol, N. J. (2009). Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO*, *20*(7), 1223–1229. https://doi.org/10.1093/annonc/mdn786 [doi]

Conroy, T., Hammel, P., Hebbar, M., Ben Abdelghani, M., Wei, A. C., Raoul, J.-L., … Bachet, J.-B. (2018). FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. *New England Journal of Medicine*, *379*(25), 2395–2406. https://doi.org/10.1056/nejmoa1809775

Coughlin, S. S., Calle, E. E., Patel, A. V., & Thun, M. J. (2000). Predictors of pancreatic cancer mortality among a large cohort of United States adults. *Cancer Causes and Control*, *11*(10), 915–923. https://doi.org/10.1023/A:1026580131793

Cristofanilli, M., Budd, G. T., Ellis, M. J., Stopeck, A., Matera, J., Miller, M. C., … Hayes, D. F. (2004). Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *The New England Journal of Medicine*, *351*(8), 781–791. https://doi.org/10.1056/NEJMoa040766 [doi]

Cubilla, A. L., & Fitzgerald, P. J. (1976). *Morphological Lesions Associated with Human Primary Invasive Nonendocrine Pancreas Cancer1*. *36*(July), 2690–2698.

de Wilde, R., Hruban, R., Maitra, A., & Offerhaus, G. (2011). Reporting precursors to invasive pancreatic cancer : pancreatic intraepithelial neoplasia , intraductal neoplasms and mucinous cystic neoplasm. *Diagnostic Histopathology*, *18*(1), 17–30. https://doi.org/10.1016/j.mpdhp.2011.10.012

Demel, U., Tilz, G. P., Foeldes-Papp, Z., Gutierrez, B., Albert, W. H., & Bocher, O. (2004). Detection of tumour cells in the peripheral blood of patients with breast cancer. Development of a new sensitive and specific immunomolecular assay. *Journal of Experimental & Clinical Cancer Research : CR*, *23*(3), 465–468.

Den, R. B., Yousefi, K., Trabulsi, E. J., Abdollah, F., Choeurng, V., Feng, F. Y., … Karnes, R. J. (2015). Genomic classifier identifies men with adverse pathology after radical prostatectomy who benefit from adjuvant radiation therapy. *Journal of Clinical Oncology*, *33*(8), 944–951. https://doi.org/10.1200/JCO.2014.59.0026

Denzel, S., Maetzel, D., Mack, B., Eggert, C., Bärr, G., & Gires, O. (2009). *Initial activation of EpCAM cleavage via cell-to-cell contact*. *14*, 1–14. https://doi.org/10.1186/1471-2407-9-402

Dijkstra, J. R., Tops, B. B. J., Nagtegaal, I. D., van Krieken, J. H. J. M., & Ligtenberg, M. J. L. (2015). The homogeneous mutation status of a 22 gene panel justifies the use of serial sections of colorectal cancer tissue for external quality assessment. *Virchows Archiv*, *467*(3), 273–278. https://doi.org/10.1007/s00428-015-1789-5

Dotan, E., Alpaugh, R. K., Ruth, K., Benjamin, P., Denlinger, C. S., Hall, M. J., … Cohen, S. J. (2016). *Prognostic Significance of MUC-1 in Circulating Tumor Cells in Patients With Metastatic Pancreatic Adenocarcinoma*. *00*(00), 1–5.

Ducreux, M., Cuhna, A. S., Caramella, C., Hollebecque, A., Burtin, P., Goéré, D., … ESMO Guidelines Committee. (2015). Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, *26 Suppl 5*(Supplement 5), v56-68. https://doi.org/10.1093/annonc/mdv295

Elhanafi, S., Mahmud, N., Vergara, N., Kochman, M. L., Das, K. K., Ginsberg, G. G., … Chandrasekhara, V. (2018). Comparison of endoscopic ultrasound tissue acquisition methods for genomic analysis of pancreatic cancer. *Journal of Gastroenterology and Hepatology (Australia)*, (Irb 829027), 1–7. https://doi.org/10.1111/jgh.14540

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., … Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, *136*(5), E359-86. https://doi.org/10.1002/ijc.29210

Fidler, I. J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *NATURE REVIEWS*, *3*(June), 1–6.

Franco, J., Witkiewicz, A. K., & Knudsen, E. S. (2014). CDK4/6 inhibitors have potent

activity in combination with pathway selective therapeutic agents in models of pancreatic cancer. *Oncotarget*, *5*(15), 6512–6525. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4171647&tool=pmcentrez&rendertype=abstract

Freeman ML, DiSario JA, Nelson DB, et al. (2001). Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. *Gastrointest Endosc*, *54*, 425–434.

Fuccio, L., & Larghi, A. (2014). Endoscopic ultrasound-guided fine needle aspiration: How to obtain a core biopsy? *Endoscopic Ultrasound*, *3*(2), 71–81. https://doi.org/10.4103/2303-9027.123011

Fukushima, N., Walter, K. M., Ueki, T., Sato, N., Matsubayashi, H., Cameron, J. L., … No, C. (2003). *Diagnosing Pancreatic Cancer Using Methylation Specific PCR Analysis of Pancreatic Juice nd sc ie nc e . ot fo r d is t rib ut n* . (February), 78–83.

Furukawa, T., Chiba, R., Kobari, M., Matsuno, S., Nagura, H., & Takahashi, T. (1994). Varying grades of epithelial atypia in the pancreatic ducts of humans. Classification based on morphometry and multivariate analysis and correlated with positive reactions of carcinoembryonic antigen. *Archives of Pathology & Laboratory Medicine*, *118*(3), 227–234.

Gerlinger, M., Andrew, M., Horswell, S., Larkin, J., Endesfelder, D., Math, D., … Ph, D. (2012). Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *NEJM*, *366*(10), 883–892.

Gertler, R., Rosenberg, R., Fuehrer, F., Dahm, M., Nekarda, H., & Siewert, J. (2003a). Detection of circulating tumor cells in blood using an optimized density gradient centrifugation. *Recent Results Cancer Res.*, *162*, 149–155.

Gertler, R., Rosenberg, R., Fuehrer, K., Dahm, M., Nekarda, H., & Siewert, J. R. (2003b). Detection of circulating tumor cells in blood using an optimized density gradient centrifugation. *Recent Results in Cancer Research.Fortschritte Der Krebsforschung.Progres Dans Les Recherches Sur Le Cancer*, *162*, 149–155.

Gorges, T. M., Tinhofer, I., Drosch, M., Rose, L., Zollner, T. M., Krahn, T., & von Ahsen, O. (2012). Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition. *BMC Cancer*, *12*, 178. https://doi.org/10.1186/1471-2407-12-178 [doi]

Gozgit, J. M., Wong, M. J., Moran, L., Wardwell, S., Mohemmad, Q. K., Narasimhan, N. I., … Rivera, V. M. (2012). Ponatinib (AP24534), a Multitargeted Pan-FGFR Inhibitor with Activity in Multiple FGFR-Amplified or Mutated Cancer Models.

Molecular Cancer Therapeutics, *11*(3), 690–699. https://doi.org/10.1158/1535-7163.MCT-11-0450

Greaves, M., & Maley, C. C. (2012a). Clonal evolution in cancer. *Nature*, *481*(7381), 306–313. https://doi.org/10.1038/nature10762

Greaves, M., & Maley, C. C. (2012b). Clonal evolution in cancer. *Nature*, *481*(7381), 306–313. https://doi.org/10.1038/nature10762

Greenhalf, W., Grocock, C., Harcus, M., & Neoptolemos, J. (2009). Screening of high-risk families for pancreatic cancer. *Pancreatology : Official Journal of the International Association of Pancreatology (IAP) ... [et Al.]*, *9*(3), 215–222. https://doi.org/10.1159/000210262

Griwatz, C., Brandt, B., Assmann, G., & Zänker, K. S. (1995). An immunological enrichment method for epithelial cells from peripheral blood. *Journal of Immunological Methods*, *183*(2), 251–265. https://doi.org/10.1016/0022-1759(95)00063-G

Grocock, C., Vitone, L., Harcus, M., Neoptolemos, J., Raraty, M., & Greenhalf, W. (2007). *Familial Pancreatic Cancer : a review and latest advances*. *52*, 37–49.

Gupta, G. P., & Massagué, J. (2006). Cancer Metastasis: Building a Framework. *Cell*, *127*(4), 679–695. https://doi.org/10.1016/j.cell.2006.11.001

Hahn, S. A., Greenhalf, B., Ellis, I., Sina-frey, M., Rieder, H., Korte, B., … Bartsch, D. K. (2003). *BRCA2 Germline Mutations in Familial Pancreatic Carcinoma*. *95*(3).

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, *100*, 57–70. https://doi.org/10.1007/s00262-010-0968-0

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Harouaka, R. A., Nisic, M., & Zheng, S. Y. (2013). Circulating tumor cell enrichment based on physical properties. *Journal of Laboratory Automation*, *18*(6), 455–468. https://doi.org/10.1177/2211068213494391 [doi]

Heitzer, E., Ulz, P., & Geigl, J. B. (2015). Circulating Tumor DNA as a Liquid Biopsy for Cancer. *Clinical Chemistry*, *61*(1), 112–123.

Hingorani, S. R., Iii, E. F. P., Maitra, A., Rajapakse, V., King, C., Jacobetz, M. A., … Tuveson, D. A. (2003). *Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse*. *4*(December), 437–450.

Holdenrieder, S., Stieber, P., Bodenmüller, H., Busch, M., Fertig, G., Fürst, H., … Seidel, D. (2001). Nucleosomes in serum of patients with benign and malignant diseases. *International Journal of Cancer*. *Journal International Du Cancer*, *95*(2),

114–120. https://doi.org/10.1002/1097-0215(20010320)95:2<114::AID-IJC1020>3.0.CO;2-Q

Hou, J. M., Greystoke, A., Lancashire, L., Cummings, J., Ward, T., Board, R., … Blackhall, F. H. (2009). Evaluation of circulating tumor cells and serological cell death biomarkers in small cell lung cancer patients undergoing chemotherapy. *The American Journal of Pathology*, *175*(2), 808–816. https://doi.org/10.2353/ajpath.2009.090078 [doi]

Hruban, R., Adsay, V., Saavedra, J. A., Compton, C., Ph, D., Garrett, E. S., … Ph, D. (2001). Pancreatic Intraepithelial Neoplasia. A New Nomenclature and Classification System for Pancreatic Duct Lesions. *The American Journal of Surgical Pathology*, *25*(5), 579–586.

Hruban, R. H., Goggins, M., Parsons, J., & Kern, S. E. (2000). *Progression Model for Pancreatic Cancer Progression Model for Pancreatic Cancer 1*. *6*(August), 2969–2972.

Hudson, T. J., Anderson, W., Aretz, A., Barker, A. D., Bell, C., Bernabé, R. R., … Yang, H. (2010). International network of cancer genome projects. *Nature*, *464*(7291), 993–998. https://doi.org/10.1038/nature08987

Ito, D., Fujimoto, K., Mori, T., Kami, K., Koizumi, M., Toyoda, E., … Doi, R. (2006). In vivo antitumor effect of the mTOR inhibitor CCI-779 and gemcitabine in xenograft models of human pancreatic cancer. *International Journal of Cancer. Journal International Du Cancer*, *118*(9), 2337–2343. https://doi.org/10.1002/ijc.21532

Ito, S., Nakanishi, H., Hirai, T., Kato, T., Kodera, Y., Feng, Z., … Tatematsu, M. (2002). Quantitative detection of CEA expressing free tumor cells in the peripheral blood of colorectal cancer patients during surgery with real-time RT-PCR on a LightCycler. *Cancer Letters*, *183*(2), 195–203. https://doi.org/S030438350200157X [pii]

Jang, J. Y., Tada, M., Salvia, R., Levy, P., Shimizu, Y., Wolfgang, C. L., … Fernández-del Castillo, C. (2017). Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatology*, *17*(5), 738–753. https://doi.org/10.1016/j.pan.2017.07.007

Juratli, M. A., Sarimollaoglu, M., Siegel, E. R., Nedosekin, D. A., Galanzha, E. I., Suen, J. Y., & Zharov, V. P. (2014). Real-time monitoring of circulating tumor cell release during tumor manipulation using in vivo photoacoustic and fluorescent flow cytometry. *Head and Neck-Journal for the Sciences and Specialties of the*

Head and Neck, *36*(8), 1207–1215. https://doi.org/10.1002/hed.23439

Kalluri, R., & Weinberg, R. a. (2009). The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation*, *119*(6), 1420–1428. https://doi.org/10.1172/JCI39104.1420

Kanda, M., Knight, S., Topazian, M., Syngal, S., Farrell, J., Lee, J., … Goggins, M. (2012). Mutant GNAS detected in duodenal collections of secretin-stimulated pancreatic juice indicates the presence or emergence of pancreatic cysts. *Gut*, 1024–1033. https://doi.org/10.1136/gutjnl-2012-302823

Kanda, Mitsuro, Knight, S., Topazian, M., Syngal, S., Farrell, J., Lee, J., … Goggins, M. (2013). Mutant GNAS detected in duodenal collections of secretin-stimulated pancreatic juice indicates the presence or emergence of pancreatic cysts. *Gut*, *62*(7), 1024–1033. https://doi.org/10.1136/gutjnl-2012-302823

Kanda, Mitsuro, Sadakari, Y., Borges, M., Topazian, M., Farrell, J., Syngal, S., … Goggins, M. (2013). Mutant TP53 in duodenal samples of pancreatic juice from patients with pancreatic cancer or high-grade dysplasia. *Clinical Gastroenterology and Hepatology*, *11*(6), 719-730.e5. https://doi.org/10.1016/j.cgh.2012.11.016

Kasimir-Bauer, S., Hoffmann, O., Wallwiener, D., Kimmig, R., & Fehm, T. (2012). Expression of stem cell and epithelial-mesenchymal transition markers in primary breast cancer patients with circulating tumor cells. *Breast Cancer Research : BCR*, *14*(1), R15. https://doi.org/bcr3099 [pii]

Khoja, L., Backen, A., Sloane, R., Menasce, L., Ryder, D., Krebs, M., … Dive, C. (2012). A pilot study to explore circulating tumour cells in pancreatic cancer as a novel biomarker. *British Journal of Cancer*, *106*(3), 508–516. https://doi.org/10.1038/bjc.2011.545 [doi]

Kim, E. S., Herbst, R. S., Wistuba, I. I., Jack Lee, J., Blumenschein, G. R., Tsao, A., … Hong, W. K. (2011). The BATTLE trial: Personalizing Therapy for Lung Cancer. *Cancer Discovery*, *1*(1), 44–53. https://doi.org/10.1158/2159-8274.CD-10-0010

Kocarnik, J. M., Shiovitz, S., & Phipps, A. I. (2015). Molecular phenotypes of colorectal cancer and potential clinical applications. *Gastroenterology Report*, *3*(September), 1–8. https://doi.org/10.1093/gastro/gov046

Korolev, K. S., Xavier, J. B., & Gore, J. (2014). Turning ecology and evolution against cancer. *Nature Publishing Group*, *14*(5), 371–380. https://doi.org/10.1038/nrc3712

Larsson, S. C., Permert, J., Håkansson, N., Näslund, I., Bergkvist, L., & Wolk, a. (2005). Overall obesity, abdominal adiposity, diabetes and cigarette smoking in relation to the risk of pancreatic cancer in two Swedish population-based cohorts.

British Journal of Cancer, *93*(11), 1310–1315. https://doi.org/10.1038/sj.bjc.6602868

Lj, V., Greenhalf, W., Nr, H., & Jp, N. (2005). *Hereditary pancreatitis and secondary screening for early pancreatic cancer*. *50*.

Lüttges, J., Schlehe, B., Menke, M. a. O. H., Vogel, I., Henne- Bruns, D., & Klöppel, G. (1999). The K- ras mutation pattern in pancreatic ductal adenocarcinoma usually is identical to that in associated normal, hyperplastic, and metaplastic ductal epithelium. *Cancer*, *85*(8), 1703–1710. https://doi.org/10.1002/(SICI)1097-0142(19990415)85:8<1703::AID-CNCR9>3.3.CO;2-I

Mahadevan, D., & Von Hoff, D. D. (2007). Tumor-stroma interactions in pancreatic ductal adenocarcinoma. *Molecular Cancer Therapeutics*, *6*(4), 1186–1197. https://doi.org/10.1158/1535-7163.MCT-06-0686

Maley, C. C., Galipeau, P. C., Finley, J. C., Wongsurawat, V. J., Li, X., Sanchez, C. A., … Reid, B. J. (2006). *Genetic clonal diversity predicts progression to esophageal adenocarcinoma*. *38*(4), 468–473. https://doi.org/10.1038/ng1768

Malka, D. (2002). Risk of pancreatic adenocarcinoma in chronic pancreatitis. *Gut*, *51*(6), 849–852. https://doi.org/10.1136/gut.51.6.849

Mani, S. a., Guo, W., Liao, M.-J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., … Weinberg, R. a. (2008). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell*, *133*(4), 704–715. https://doi.org/10.1016/j.cell.2008.03.027

Martinez, P., Birkbak, N. J., Gerlinger, M., McGranahan, N., Burrell, R. A., Rowan, A. J., … Swanton, C. (2013). Parallel evolution of tumour subclones mimics diversity between tumours. *Journal of Pathology*, *230*(4), 356–364. https://doi.org/10.1002/path.4214

McKenzie SB. (1996). *Textbook of hematology*. Baltimore: Williams and Wilkins.

Meng, S., Tripathy, D., Frenkel, E. P., Shete, S., Naftalis, E. Z., Huth, J. F., … Uhr, J. W. (2004). Circulating tumor cells in patients with breast cancer dormancy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *10*(24), 8152–8162. https://doi.org/10/24/8152 [pii]

Michaud, D. S. (2004). *Epidemiology of pancreatic cancer*. *59*, 99–111.

Milne, R. L., Greenhalf, W., Murta-Nascimento, C., Real, F. X., & Malats, N. (2009). The inherited genetic component of sporadic pancreatic adenocarcinoma. *Pancreatology : Official Journal of the International Association of Pancreatology (IAP) ... [et Al.]*, *9*(3), 206–214. https://doi.org/10.1159/000210261

Moore, P. S., Sipos, B., Orlandini, S., Sorio, C., Real, F. X., Lemoine, N. R., … Scarpa, a. (2001). Genetic profile of 22 pancreatic carcinoma cell lines. Analysis of K-ras, p53, p16 and DPC4/Smad4. *Virchows Archiv : An International Journal of Pathology*, *439*, 798–802. https://doi.org/10.1007/s004280100474

Mosaluk, C., Hruban, R. H., & Kern, S. E. (1997). p16 and K-ras Gene Mutations Adenocarcinoma ' in the Intraductal Precursors of Human Pancreatic. *Cancer Research*, *57*, 2140–2144.

Muller, V., Stahmann, N., Riethdorf, S., Rau, T., Zabel, T., Goetz, A., … Pantel, K. (2005). Circulating TumorCells in BreastCancer:Correlation to Bone Marrow Micrometastases, Heterogeneous Response to Systemic Therapy and Low Proliferative Activity. *Clinical Cancer Research*, *11*(5), 3678–3686.

Muscat, E., Steilman, D., & Wynder, E. L. (n.d.). *Smoking and Pancreatic Cancer in Men and Women '*.

Nagrath, S., Jack, R. M., Sahai, V., & Simeone, D. M. (2016). REVIEWS IN BASIC AND CLINICAL GASTROENTEROLOGY Opportunities and Challenges for Pancreatic Circulating. *Gastroenterology*, *151*(3), 412–426. https://doi.org/10.1053/j.gastro.2016.05.052

Nagrath, S., Sequist, L. V, Maheswaran, S., Bell, D. W., Ryan, P., Balis, U. J., … Haber, D. A. (2011). Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature*, *450*(7173), 1235–1239. https://doi.org/10.1038/nature06385.Isolation

Naume, B., Borgen, E., Tossvik, S., Pavlak, N., Oates, D., & Nesland, J. M. (2004). Detection of isolated tumor cells in peripheral blood and in BM: evaluation of a new enrichment method. *Cytotherapy*, *6*(3), 244–252. https://doi.org/10.1080/14653240410006086 [doi]

Neoptolemos, J. P., Kleeff, J., Michl, P., Costello, E., Greenhalf, W., & Palmer, D. H. (2018). Therapeutic developments in pancreatic cancer: Current and future perspectives. *Nature Reviews Gastroenterology and Hepatology*, *15*(6), 333–348. https://doi.org/10.1038/s41575-018-0005-x

Nicholson, J. A., Greenhalf, W., Jackson, R., Trevor, F., Hanna, T., Harrison, S., … Evans, J. C. (2015). Incidence of Post-ERCP Pancreatitis From Direct Pancreatic Juice Collection in Hereditary Pancreatitis and Familial Pancreatic Cancer Before and After the Introduction of Prophylactic Pancreatic Stents and Rectal Diclofenac. *Pancreas*, *44*(2), 260–265.

Nies, H., P, C., A, R., I, I., Y, Z., S, K., … CJ, B. (2014). Side population cells of

pancreatic cancer show characteristics of cancer stem cells responsible for resistance and metastasis. *Target Oncology*.

Nones, K., Waddell, N., Song, S., Patch, A. M., Miller, D., Johns, A., … Grimmond, S. M. (2014). Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. *International Journal of Cancer*, *135*(5), 1110–1118. https://doi.org/10.1002/ijc.28765

Olsson, E., Winter, C., George, A., Chen, Y., Howlin, J., Tang, M. E., … Saal, L. H. (2015). Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. *EMBO Molecular Medicine*, *7*(8), 1–15. https://doi.org/10.15252/emmm.201404913

Pao, W., & Chmielecki, J. (2010). Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat Rev Cancer*, *10*(11), 760–774. Retrieved from http://dx.doi.org/10.1038/nrc2947

Papavasiliou, P., Fisher, T., Kuhn, J., Nemunaitis, J., & Lamont, J. (2010). *Circulating tumor cells in patients undergoing surgery for hepatic metastases from colorectal cancer*. *20*(1), 11–14.

Park, S. Y., Choi, G.-S., Park, J. S., Kim, H. J., Ryuk, J.-P., & Choi, W.-H. (2012). Influence of surgical manipulation and surgical modality on the molecular detection of circulating tumor cells from colorectal cancer. *Journal of the Korean Surgical Society*, *82*(6), 356–364. https://doi.org/10.4174/jkss.2012.82.6.356

Perez-Mancera, P. a., Guerra, C., Barbacid, M., & Tuveson, D. a. (2012). What We Have Learned About Pancreatic Cancer From Mouse Models. *Gastroenterology*, *142*(5), 1079–1092. https://doi.org/10.1053/j.gastro.2012.03.002

Perez-Ramirez, Canadas-Garre, Jimenez-Varo, Faus-Dader, & Calleja-Harnandez. (2015). Pharmacogenomics MET : a new promising biomarker in. *Pharmacogenomics*, *719*, 631–647.

Pesta, M., Fichtl, J., Kulda, V., Topolcan, O., & Treska, V. (2013). Monitoring of circulating tumor cells in patients undergoing surgery for hepatic metastases from colorectal cancer. *Anticancer Research*, *33*(5), 2239–2243. https://doi.org/33/5/2239 [pii]

Pisetsky, D. S. (2004). The immune response to cell death in SLE. *Autoimmunity Reviews*, *3*(7–8), 500–504. https://doi.org/10.1016/j.autrev.2004.07.010

Polzer, B., Medoro, G., Pasch, S., Fontana, F., Zorzino, L., Pestka, A., … Klein, C. a. (2014). Molecular profiling of single circulating tumor cells with diagnostic

intention. *EMBO Molecular Medicine*, *6*(11), 1371–1386. https://doi.org/10.15252/emmm.201404033

Raimondi, C., Nicolazzo, C., Gradilone, A., Molecolare, D. M., & Università, S. (2015). *Circulating tumor cells isolation : the " post-EpCAM era ." 27*(5), 461–470. https://doi.org/10.3978/j.issn.1000-9604.2015.06.02

Raty S, Sand J, Pulkkinen M, et al. (2001). Post-ERCP pancreatitis: reduction by routine antibiotics. *J Gastrointest Surg*, *5*, 339–345.

Redig, a. J., & Janne, P. a. (2015). Basket Trials and the Evolution of Clinical Trial Design in an Era of Genomic Medicine. *Journal of Clinical Oncology*, *33*(9), 975–978. https://doi.org/10.1200/JCO.2014.59.8433

Rimmer, N., Greenhalf, W., & Flaherty, L. (2014). *SOP GCLPTSS105/2 Processing and storing blood for the livc3 study*.

Rivlin, N., Brosh, R., Oren, M., & Rotter, V. (2011). Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes & Cancer*, *2*(4), 466–474. https://doi.org/10.1177/1947601911408889

Robinson, J., Thorvaldsdóttir, H., W, W., Guttman, M., Lander, E., Getz, G., & MEsirov, J. (2012). Integrative Genomics Viewer James. *Nat Biotechnol*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754.Integrative

Ross, J. S., Slodkowska, E. A., Symmans, W. F., Pusztai, L., Ravdin, P. M., & Hortobagyi, G. N. (2009). The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. *The Oncologist*, *14*(4), 320–368. https://doi.org/10.1634/theoncologist.2008-0230

Satelli, A., Mitra, A., Brownlee, Z., Xia, X., Bellister, S., Overman, M. J., … Li, S. (2014). *Epithelial – Mesenchymal Transitioned Circulating Tumor Cells Capture for Detecting Tumor Progression*. 1–9. https://doi.org/10.1158/1078-0432.CCR-14-0894

Sawabata, N., Funaki, S., Hyakutake, T., & Shintani, Y. (2016). Perioperative circulating tumor cells in surgical patients with non‐small cell lung cancer : does surgical manipulation dislodge cancer cells thus allowing them to pass into the peripheral blood ? *Surgery Today*, 2–9. https://doi.org/10.1007/s00595-016-1318-4

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors - occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Schramm, A., Friedl, T. W. P., Schochter, F., Scholz, C., de Gregorio, N., Huober, J., … Fehm, T. (2015). Therapeutic intervention based on circulating tumor cell

phenotype in metastatic breast cancer: concept of the DETECT study program. *Archives of Gynecology and Obstetrics*. https://doi.org/10.1007/s00404-015-3879-7

Shah, S. M., Ribeiro, A., Levi, J., Jorda, M., Rocha-Lima, C., Sleeman, D., … Barkin, J. (2008). EUS-guided fine needle aspiration with and without trucut biopsy of pancreatic masses. *JOP : Journal of the Pancreas*, *9*(4), 422–430. https://doi.org/v09i04a08 [pii]

Shaw, V., Bullcok, K., & Greenhalf, W. (2016). Single-Nucleotide Polymorphism to Associate Cancer Risk. *Methods in Molecular Biology (Clifton, N.J.)*, *1381*, 93–110.

Sheel, A. R. G., Harrison, S., Sarantitis, I., Nicholson, J. A., Hanna, T., Grocock, C., … Greenhalf, W. (2018). Identification of Cystic Lesions by Secondary Screening of Familial Pancreatic Cancer (FPC) Kindreds Is Not Associated with the Stratified Risk of Cancer. *American Journal of Gastroenterology*, 155–164. https://doi.org/10.1038/s41395-018-0395-y

Shen, K., Luk, S., Hicks, D. F., Elman, J. S., Bohr, S., Iwamoto, Y., … Parekkadan, B. (2014). Resolving cancer–stroma interfacial signalling and interventions with micropatterned tumour–stromal assays. *Nature Communications*, *5*, 5662. https://doi.org/10.1038/ncomms6662

Sherry, S., Ward, M., & Sirotkin, K. (1999). dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res*, *8*(9), 677–679.

Shibue, T., & Weinberg, R. A. (2017). EMT, CSCs, and drug resistance: The mechanistic link and clinical implications. *Nature Reviews Clinical Oncology*, *14*(10), 611–629. https://doi.org/10.1038/nrclinonc.2017.44

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., & Mcguire, W. L. (1987). Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2lneu Oncogene. *Science*, *235*(21), 0–5.

Smerage, J. B., Barlow, W. E., Hortobagyi, G. N., Winer, E. P., Leyland-jones, B., Srkalovic, G., … Schott, A. F. (2014). *J OURNAL OF C LINICAL O NCOLOGY Circulating Tumor Cells and Response to Chemotherapy in Metastatic Breast Cancer : SWOG S0500*. 1–8. https://doi.org/10.1200/JCO.2014.56.2561

Sottoriva, A., Spiteri, I., Piccirillo, S. G. M., Touloumis, A., Collins, V. P., Marioni, J. C., … Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences of*

the United States of America, *110*(10), 4009–4014. https://doi.org/10.1073/pnas.1219747110

Speelman, A. D., Gestel, Y. R. B. M. Van, Rutten, H. J. T., & Hingh, I. H. J. T. De. (2015). Changes in gastrointestinal cancer resection rates. *Bristish Journal of Surgery*, *102*(9), 1114–1122. https://doi.org/10.1002/bjs.9862

Spencer, D. H., Sehn, J. K., Abel, H. J., Watson, M. a., Pfeifer, J. D., & Duncavage, E. J. (2013). Comparison of Clinical Targeted Next-Generation Sequence Data from Formalin-Fixed and Fresh-Frozen Tissue Specimens. *The Journal of Molecular Diagnostics*, *15*(5), 623–633. https://doi.org/10.1016/j.jmoldx.2013.05.004

Stott, S. L., Lee, R. J., Nagrath, S., Yu, M., Miyamoto, D. T., Ulkus, L., … Maheswaran, S. (2010). Isolation and Characterization of Circulating Tumor Cells from Patients with Localized and Metastatic Prostate Cancer. *Science Translational Medicine*, *2*(25), 25ra23--25ra23. https://doi.org/10.1126/scitranslmed.3000403

Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, *458*(7239), 719–724. https://doi.org/10.1038/nature07943

Stremitzer, S., Zhang, W., Yang, D., Ning, Y., Stintzing, S., Sebio, A., … Lenz, H.-J. (2015). Genetic variations in angiopoietin and pericyte pathways and clinical outcome in patients with resected colorectal liver metastases. *Cancer*, 1–8. https://doi.org/10.1002/cncr.29259

Szpechcinski, a, Chorostowska-Wynimko, J., Struniawski, R., Kupis, W., Rudzinski, P., Langfort, R., … Orlowski, T. (2015). Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. *British Journal of Cancer*, *113*(3), 476–483. https://doi.org/10.1038/bjc.2015.225

Tanaka, M., Fernández-Del Castillo, C., Adsay, V., Chari, S., Falconi, M., Jang, J. Y., … Yamao, K. (2012). International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. *Pancreatology*, *12*(3), 183–197. https://doi.org/10.1016/j.pan.2012.04.004

Tops, B. B., Normanno, N., Kurth, H., Amato, E., Mafficini, A., Rieber, N., … Laurent-Puig, P. (2015). Development of a semi-conductor sequencing-based panel for genotyping of colon and lung cancer by the Onconetwork consortium. *BMC Cancer*, *15*(1), 1–9. https://doi.org/10.1186/s12885-015-1015-5

UK, P. C. (2019). No Title. Retrieved April 20, 2019, from https://www.pancreaticcancer.org.uk/information-and-support/clinical-trials/find-a-clinical-trial/open-clinical-trials/primus-002/

Vitone, L. J., Greenhalf, W., McFaul, C. D., Ghaneh, P., & Neoptolemos, J. P. (2006). The inherited genetics of pancreatic cancer and prospects for secondary screening. *Best Practice and Research: Clinical Gastroenterology*, *20*(2), 253–283. https://doi.org/10.1016/j.bpg.2005.10.007

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nat. Med.*, *10*(8), 789–799. https://doi.org/10.1038/nm1087

von Burstin, J., Eser, S., Paul, M. C., Seidler, B., Brandl, M., Messer, M., … Saur, D. (2009). E-Cadherin Regulates Metastasis of Pancreatic Cancer In Vivo and Is Suppressed by a SNAIL/HDAC1/HDAC2 Repressor Complex. *Gastroenterology*, *137*(1), 361-371.e5. https://doi.org/10.1053/j.gastro.2009.04.004

Waddell, N., Pajic, M., Patch, A., Chang, D. K., Kassahn, K. S., Bailey, P., … Grimmond, S. M. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature*. https://doi.org/10.1038/nature14169

Wagle, N., Berger, M. F., Davis, M. J., Blumenstiel, B., DeFelice, M., Pochanard, P., … Garraway, L. a. (2012). High-Throughput Detection of Actionable Genomic Alterations in Clinical Tumor Samples by Targeted, Massively Parallel Sequencing. *Cancer Discovery*, *2*(1), 82–93. https://doi.org/10.1158/2159-8290.CD-11-0184

Wald, N. J. (2008). Guidance on terminology. *Journal of Medical Screening*, *15*(1), 50. https://doi.org/10.1258/jms.2008.008got

Weng, L., Wu, X., Gao, H., Mu, B., Li, X., Wang, J.-H., … Wu, H. (2010). MicroRNA profiling of clear cell renal cell carcinoma by whole-genome small RNA deep sequencing of paired frozen and formalin-fixed, paraffin-embedded tissue specimens. *The Journal of Pathology*, *222*(1), 41–51. https://doi.org/10.1002/path.2736

Wiersema, M. J., Levy, M. J., Harewood, G. C., Vazquez-Sequeiros, E., Jondal, M. Lou, & Wiersema, L. M. (2002). Initial experience with EUS-guided trucut needle biopsies of perigastric organs. *Gastrointestinal Endoscopy*, *56*(2), 275–278. https://doi.org/10.1016/S0016-5107(02)70193-4

Wilentz, R. E., Geradts, J., Maynard, R., Offerhaus, G. J. A., Kang, M., Goggins, M., … Hruban, R. H. (1998). *Inactivation of the p16 (INK4A) Tumor-suppressor Gene in Pancreatic Duct Lesions : Loss of Intranuclear Expression1*. *16*.

Wilentz, R. E., Iacobuzio-donahue, C. A., Argani, P., Mccarthy, D. M., Parsons, J. L., Yeo, C. J., … Hruban, R. H. (2006). *Loss of Expression of Dpc4 in Pancreatic Intraepithelial Neoplasia : Evidence That DPC4 Inactivation Occurs Late in*

Neoplastic Progression 1. (410), 2002–2006.

Wilson, J., & Jungner, G. (1968). Principles and practice of screening for disease. Retrieved April 1, 2016, from WHO website: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf

Wu, S., Liu, S., Liu, Z., Huang, J., Pu, X., Li, J., … Xu, J. (2015). Classification of circulating tumor cells by epithelial-mesenchymal transition markers. *PloS One*, *10*(4), e0123976. https://doi.org/10.1371/journal.pone.0123976

Yachida, S, & Iacobuzio-Donahue, C. a. (2013). Evolution and dynamics of pancreatic cancer progression. *Oncogene*, *32*(45), 5253–5260. https://doi.org/10.1038/onc.2013.29

Yachida, Shinichi, Jones, S., Bozic, I., Antal, T., Leary, R., Kamiyama, M., … Iacobuzio-, C. A. (2011). Distant Metastasis Occurs Late during the Genetic Evolution of Pancreatic Cancer. *Nature*, *467*(7319), 1114–1117. https://doi.org/10.1038/nature09515.Distant

Yadav, S. S., Li, J., Lavery, H. J., Yadav, K. K., & Tewari, A. K. (2015). Next-generation sequencing technology in prostate cancer diagnosis, prognosis, and personalized treatment. *Urologic Oncology: Seminars and Original Investigations*, *33*(6), 1–13. https://doi.org/10.1016/j.urolonc.2015.02.009

Yamada, S., Fuchs, B. C., Fujii, T., Shimoyama, Y., Sugimoto, H., Nomoto, S., … Nakao, A. (2013). Epithelial-to-mesenchymal transition predicts prognosis of pancreatic cancer. *Surgery (United States)*, *154*(5), 946–954. https://doi.org/10.1016/j.surg.2013.05.004

Yan, L., McFaul, C., Howes, N., Leslie, J., Lancaster, G., Wong, T., … Greenhalf, W. (2005). Molecular Analysis to Detect Pancreatic Ductal Adenocarcinoma in High-Risk Groups. *Gastroenterology*, *128*(7), 2124–2130. https://doi.org/10.1053/j.gastro.2005.03.006

Yang, C., Shi, D., Wang, S., Wei, C., Zhang, C., & Xiong, B. (2018). Prognostic value of pre-and post-operative circulating tumor cells detection in colorectal cancer patients treated with curative resection: A prospective cohort study based on iset device. *Cancer Management and Research*, *10*, 4135–4144. https://doi.org/10.2147/CMAR.S176575

Young, G., Wang, K., He, J., Otto, G., Hawryluk, M., Zwirco, Z., … Ross, J. S. (2013). Clinical next-generation sequencing successfully applied to fine-needle aspirations of pulmonary and pancreatic neoplasms. *Cancer Cytopathology*, *121*(12), 688–694. https://doi.org/10.1002/cncy.21338

Zhang, J., Liu, J., Sun, J., Chen, C., Foltz, G., & Lin, B. (2014). Identifying driver mutations from sequencing data of heterogeneous tumors in the era of personalized genome sequencing. *Briefings in Bioinformatics*, *15*(2), 244–255. https://doi.org/10.1093/bib/bbt042

Zhang, Q., Shan, F., Li, Z., Gao, J., Li, Y., Shen, L., … Lu, M. (2018). A prospective study on the changes and clinical significance of pre-operative and post-operative circulating tumor cells in resectable gastric cancer. *Journal of Translational Medicine*, *16*(1), 1–12. https://doi.org/10.1186/s12967-018-1544-1