Circulating Tumour DNA in Advanced Thyroid Cancer

David Matthew Allin *MBBS MA FRCS*

The Institute of Cancer Research

The Royal Marsden NHS Foundation Trust

MD (Res)

Contents

Detailed Contents

Declaration

I declare that the work presented herein is my own, unless otherwise stated. It was carried out under the supervision of Dr Kate Newbold, Mr Dae Kim and Professor Kevin Harrington.

David Allin *MBBS MA FRCS*

Abstract

Conventional biomarkers currently used in thyroid cancer are not disease specific and fluctuate in advanced disease, making interpretation difficult. Circulating tumour DNA (ctDNA) has been shown to be a useful biomarker in other solid tumours. The main objective of this thesis was to ascertain if ctDNA has potential as a novel biomarker in thyroid cancer. This was undertaken through examining the hypothesis that ctDNA is detectable in the plasma of patients with advanced thyroid cancer. In support of this, and to help with future studies, two further hypotheses were examined: that a thyroid-specific targeted gene panel would offer superiority over a currently used generic panel, and that resistance to targeted therapy in thyroid cancer is due to a genetic variant detectable in ctDNA.

Through multi-mutational analysis of plasma, ctDNA was detected in the majority of patients with advanced thyroid cancer. ctDNA measurement may offer superiority over conventional markers in several scenarios: earlier detection of progression in medullary thyroid cancer; as an alternative biomarker when conventional markers are not available; more rapid assessment of the disease status in response to targeted therapies, thereby potentially allowing prompter discontinuation of futile therapies.

A novel thyroid-specific targeted gene panel was designed and validated successfully. This performed favourably compared to a generic panel, with lower costs, faster workflow and greater detection rate. Therefore, it will be used in future studies.

Lastly, whole exome sequencing of plasma samples taken from patients with advanced thyroid cancer pre- and post-resistance to targeted therapies may be a useful technique for discovering genomic mechanisms of resistance. Several leads were generated for putative resistance-conferring variants, and areas in the workflow were identified that could be improved to allow for more successful future studies.

These early results support the hypothesis that ctDNA may be a clinically useful biomarker in thyroid cancer. Clinical validity and utility will need to be confirmed in further large prospective trials.

Acknowledgements

Throughout my time at the ICR, both in the labs I worked in, and in writing up my thesis, I have received a great deal of support from many people, for which I will always be grateful.

Firstly, thank you to my supervisors, Dr Kate Newbold and Mr Dae Kim, as well as Professor Harrington. This last year has been a busy one, and without their help and patience none of this would have been possible.

Thank you to everyone at the CMP molecular diagnostics laboratory for taking me under their wing. They taught me the scientific techniques underpinning this work, all of which were entirely foreign to a clinician – so thanks for opening my eyes to the world of genomics! I would especially like to thank Dr Mike Hubank for his support and guidance. Special mentions also to Ridwan Sheikh, Paul Carter, Hannah Bye and Paula Proszek. Your help and tutelage were immense, and I am so very appreciative of you donating your time to helping me.

I am also indebted to Professor Nick Turner for welcoming me into his lab at Chelsea and for his insightful comments at my transfer viva which helped point my thesis in the right direction. Further huge thanks are due to “El Jefe” Dr Isaac Garcia-Murillas, for showing me the way, and to Dr Ben O’Leary for putting up with my ignorance and never-ending questions.

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Thyroid Cancer

1.1.1 The Thyroid Gland

The thyroid is a bilobed endocrine gland. It is situated in the midline of the neck, just anterior to the upper trachea, and weighs approximately 25-30 grams in a healthy adult. Its primary function is the secretion of thyroid hormones. Two of these hormones are *tri-iodothyronine* (T3) and *thyroxine* (T4) that have a myriad of physiological effects, but in particular act to increase basal metabolic rate and affect protein synthesis. They are also essential for normal neurodevelopment. The thyroid also secretes a third hormone, *calcitonin*, whose function is poorly understood in humans but may play a minor role in the homeostatic control of blood calcium concentration.

1.1.2 Overview of thyroid histology

The microscopic architecture of thyroid tissue is made up of follicles and surrounding cells (**Figure 1-1**). Each follicle is filled with colloid that contains thyroid hormone precursor proteins, mainly thyroglobulin (Tg). The colloid is surrounded by a single layer of epithelial follicular cells, which synthesise T3 and T4 through iodination and subsequent cleavage of thyroglobulin. This requires a significant uptake of iodine by follicular cells, which is actively transported into the cells via the sodium-iodide symporter (NIS). This physiological feature of follicular cells is exploited in the treatment of cancers derived from follicular cells, through the use of radio-iodine therapy. Parafollicular cells are found between these follicles. Also known as C-cells, these secrete calcitonin.

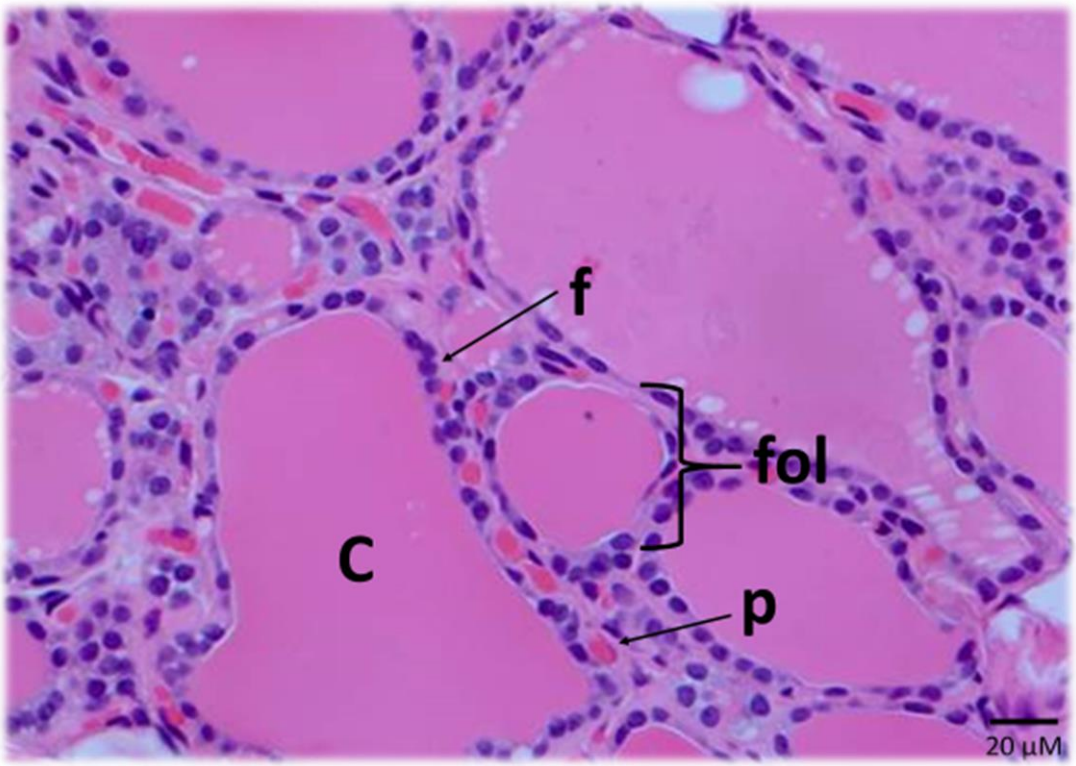

Figure 1-1. *Microscopic thyroid tissue architecture. C: colloid, fol: follicle, f: follicular cell, p: parafollicular cell*

1.1.3 Cancer genesis

The stepwise accumulation of a variety of genetic or epigenetic aberrations lead to the transformation of benign cells to a malignant state. These alterations can cause activating or inactivating mutations, changes in gene expression patterns, microRNA dysregulation and abnormal gene methylation. In thyroid cancer, the molecular alterations leading to these effects are due mainly to either small changes in bases, such as single nucleotide variations (SNVs) and small insertions/deletions (indels), or chromosomal re-arrangements such as fusions. These events usually occur spontaneously in somatic cells, leading to sporadic cases. Alternatively, mutations can involve germ-cells allowing defects to be passed from parent to offspring, leading to hereditary forms of cancer.

1.1.4 Histological subtypes of thyroid carcinoma

Traditionally, thyroid malignancies have been categorised according to histological appearance. The majority of thyroid cancers arise from follicular cells. These account for around 90% of thyroid malignancies and are termed differentiated thyroid cancer (DTC). Most DTCs are classed as *papillary*: described as malignant epithelial cells showing follicular cell differentiation that have characteristic *papillae.* These represent 80-85% of thyroid cancers (1). A further 5% are follicular and these are differentiated thyroid cancers lacking the diagnostic features of PTC (2). Conventional first line management of DTCs includes surgical excision followed by consideration of radio-iodine ablation (3).

Other, less differentiated, subtypes of thyroid cancer arising from follicular cells include poorly-differentiated thyroid carcinoma (PDTC) and anaplastic, or undifferentiated, thyroid carcinoma (ATC). Both PDTC and ATC are significantly more aggressive and have worse outcomes than other types. They usually arise from progressive de-differentiation of papillary and follicular carcinomas, although ATCs may occur *de novo* (4).

Medullary cancers arise from the parafollicular cells and account for approximately 3-5% of thyroid cancers. MTC has fewer treatment options than DTC as parafollicular cells do not take up iodine, therefore radioiodine is not a useful treatment modality. MTC has less good outcomes in terms of prognosis compared with PTC and FTC. A summary of thyroid cancer subtype frequency is shown in Table 1-1**.**

SUBTYPE	*FREQ (%)*
PAPILLARY	83
FOLLICULAR	5
MEDULLARY	3
POORLY DIFF.	2
ANAPLASTIC	1

Table 1-1. *Approximate thyroid histology subtype frequencies. Poorly diff: poorly differentiated carcinomas. Data from Aschebrook-Kilfoy et al. (5) and Sobrinho-SImoes et al. (2)*

1.1.5 Molecular pathology of thyroid cancer subtypes

1.1.5.1 Overview of MAPK and PI3K pathways

Two signalling pathways are frequently involved in the carcinogenesis of thyroid cancers: the Mitogen activated protein kinase (MAPK) and Phosphoinositide 3-kinase (PI3K) pathways (**Figure 1-2**). Both function to allow transmission of signals from a variety of cell membrane related tyrosine kinase receptors to the nucleus.

The MAPK pathway is a critical cellular signal transduction pathway involved in the regulation of cell cycle entry and proliferation. It communicates extracellular mitogenic signals, such as epidermal growth factor, to the nucleus leading to the altered expression of a variety of genes. It does this through a receptor tyrosine kinase, such as RET or NTRK1, that triggers the stepwise activation of a cascade of several protein kinases. This signalling mechanism plays an important role in the regulation of cell proliferation, differentiation and apoptosis. Disruption of the MAPK pathway appears to be a crucial step in the pathogenesis of many cancers.

Similarly, the PI3K pathway is also involved in cell cycle regulation. In this pathway the receptor tyrosine kinase activates Phosphoinositide 3-kinase (PI3K) which phosphorylates Protein kinase B (AKT). AKT is a protein kinase with a number of downstream effectors involving activated mammalian target of rapamycin (mTOR) that affect cell growth and proliferation.

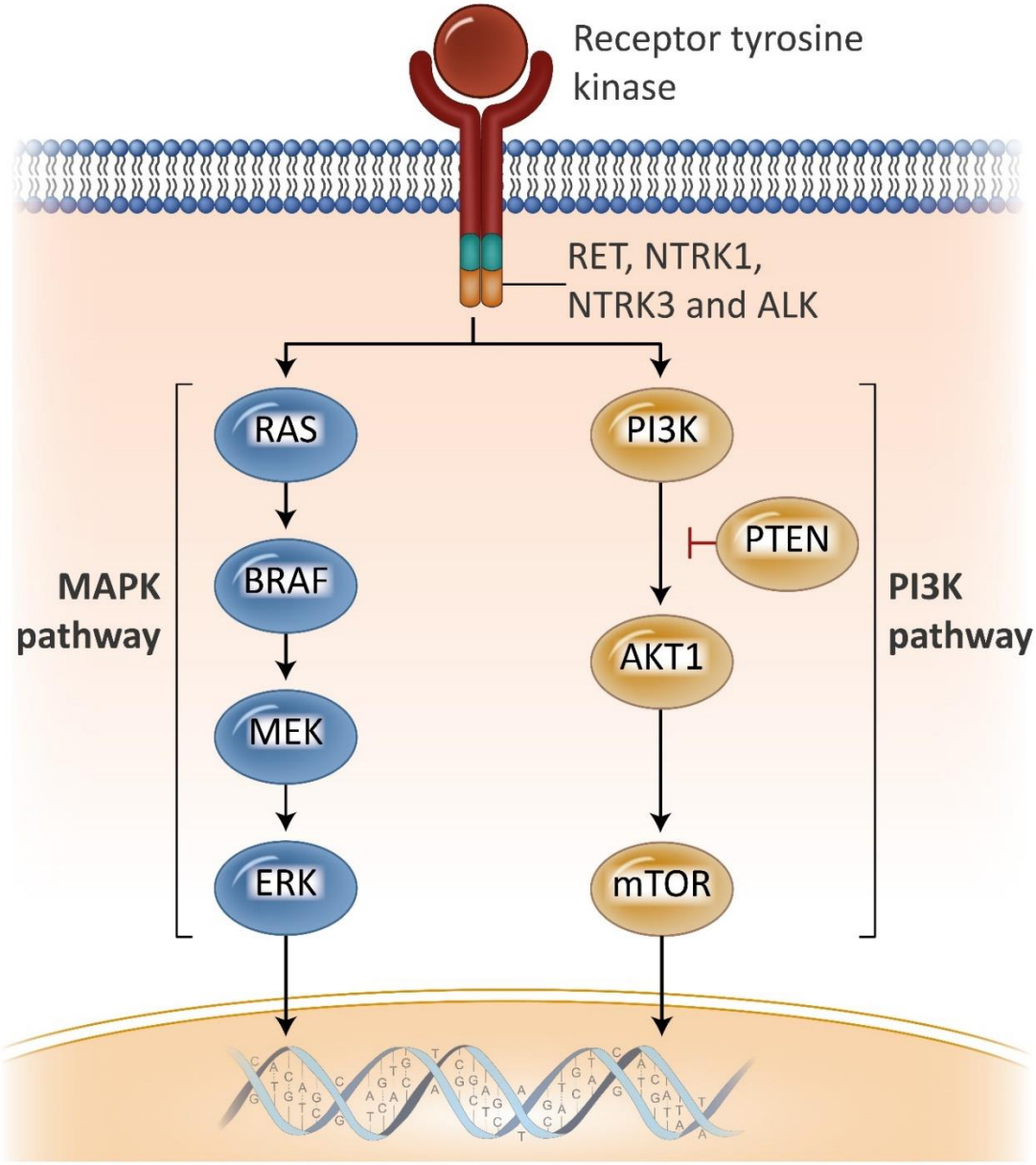

Figure 1-2. Simplified representation of MAPK and PI3K pathways

1.1.5.2 Common driver mutations associated with thyroid malignancy

RAS oncogene

The Rat Sarcoma (*RAS*) genes encode for a subfamily of GTPases that link signalling from cell membrane receptors and G-protein coupled receptors to several signalling cascades, including the MAPK and PI3K pathways. Three types have been described in humans: *HRAS* (**H**arvey sarcoma virus), *KRAS* (**K**irsten sarcoma virus) and *NRAS* (**N**euroblastoma) (6). *RAS* activating mutations are some of the most common in all cancer biology. Variants found in thyroid cancer are usually missense mutations affecting codon 61 of *NRAS* or *HRAS* and, less frequently, in codons 12 and 13. These mutations cause the loss of the GTPase activity, leading to constitutive activation of the kinase. *RAS* mutations are the most commonly found in follicular cancers (7)

BRAF oncogene

BRAF is a member of the Rapidly Accelerated Fibrosarcoma (RAF) kinases, of which three are described (A-RAF, B-RAF, C-RAF). This kinase is activated by the RAS protein, and in turn activates MAPK/ERK kinase (MEK) by phosphorylation. *BRAF* is a well-known oncogene and is recognised as a key driver mutation in several cancers, including DTC. Activating mutations allow MAPK signal transduction without requiring RAS activity. The vast majority of *BRAF* mutations (95%) are a single base substitution at coding base 1799 of Thymine to Adenine (c.1799A>T). This leads to a missense mutation, resulting in the codon at amino acid position 600 changing from Valine to Glutamic acid (p.V600E) (8). *BRAF* mutations are the most common driver mutations in thyroid cancer, present in up to 74% of papillary tumours (9). They are associated with poorer clinical prognosis in PTC (10) and a large meta-analysis has

demonstrated that the *BRAF* c.1799C>T variant is significantly associated with increased tumour size, extra-thyroid extension and clinical stage (9). The effect on survival, however, from a *BRAF* mutation alone is relatively modest: the associated hazard ratio in PTC is 1.16. Yet the combination of *BRAF* V600E with a *TERT* variant (described below) has a more profound effect on mortality, and markedly increases the hazard ratio in to 8.71 (11) and so it would appear that both variants act synergistically.

PTEN tumour suppressor gene

Phosphatase and TENsin homolog (PTEN) is an inhibitor of the PI3K pathway and acts as a tumour suppressor gene. It inhibits this pathway through dephosphorylation, thereby reducing the activity of PI3K. Inactivating variants therefore prevent negative feedback and reduce control of the PI3K pathway. Germline mutations in *PTEN* lead to Cowden Syndrome, of which increased rates of thyroid malignancy are a feature.

RET oncogene

The Re-arranged during Transfection (*RET*) gene codes for one of the tyrosine kinase receptors involved in the MAPK pathway. *RET* is normally only expressed in parafollicular thyroid cells and not in follicular cells. Mutations in this gene are rare in DTC but occur frequently in MTC. Single nucleotide variations in *RET* are particularly associated with MTC. Germline mutations in *RET* are found in familial MTC and Multiple Endocrine Neoplasia (MEN) Type 2.

RET is also recurrently involved in the creation of fusion oncoproteins in thyroid cancer, particularly in PTC. In these fusions the 3' RET receptor tyrosine kinase domain fuses with the 5' active promoter region of another other gene, which leads to chronic activation of the MAPK pathway (12, 13). *RET* fusions usually form chimeras with genes that lie on the long (q) arm of chromosome 10 and are thus in close spatial proximity to it, for example *PTC1* (CCDC6) or *PTC3* (NCOA4).

Chromosomal rearrangements have a strong association with ionising radiation. These types of genomic aberrations in thyroid cancers are seen at higher frequencies in tumours of patients exposed to therapeutic radiation or accidental radiation exposure such as those seen in the survivors of the Chernobyl nuclear power plant disaster and atomic bombings in Japan during the Second World War (14, 15).

TERT oncogene

Telomeres are regions of non-coding, repetitive nucleotide sequences present at each end of chromosomes. During cell division, DNA duplication does not occur at the ends of the chromosomes and so a small fragment of DNA is lost during each cycle. Telomeres act as a buffer to protect against loss of coding DNA. The Telomerase Reverse Transcriptase *(TERT)* gene encodes a reverse transcriptase subunit of the enzyme telomerase. Telomerase maintains the telomere length at the end of chromosomes.

Mutations in *TERT* usually involve the promoter region of the gene, at either 126 or 146 bp upstream from the start codon. These variants lead to new binding motifs for transcription factors leading to inappropriate activation of *TERT* (16). This likely aids in achieving cellular immortalisation through increased telomerase activity and thereby increases the likelihood of further oncogenic events. Thyroid malignancies with *TERT* mutations behave more aggressively (17), and lead to increased mortality, with a hazard ratio as high as 6.00 (11) in PTC. As mentioned previously, this effect is further compounded when associated with *BRAF* variants (11, 18).

Pax8/PPARG fusion protein

Paired Box 8 (*Pax8*) codes for a transcription factor required for normal thyroid tissue development. In mature thyroid tissue it induces the expression of various thyroid-specific genes. Peroxisome proliferator-activated receptor gamma (PPARG) is a nuclear transcription factor involved in glucose and lipid metabolism in muscle and adipose tissue that is not usually expressed in thyroid tissues (19). Fusion between these two genes results in over expression of a Pax8/PPARG fusion protein (PPFP), found in up to 30% of follicular cancers. As yet, it is unknown how PPFP exerts its carcinogenic effect.

1.1.5.3 Papillary Thyroid Cancer

Papillary thyroid cancers are now thought to encompass more than one pathological entity. Several subtypes have been suggested, defined by a different driver mutation (20) that affect the MAPK pathway. Three types of genomic aberrations account for the majority of these,

with *BRAF* V600E variants being the most common (21). These are summarised in **Table 1-2**. In the remaining 13% there does not appear to be a recurrent genomic event allowing simple categorisation.

Variant type	Associated PTC sub-type	Freq (%)
***BRAF* p.V600E**	Classical *or* Tall-cell	60
***RAS* variants**	FvPTC	15
Chromosomal rearrangement*	Classical	12

Table 1-2. Frequency of histological subtypes of PTC according to driver mutation. **Chromosomal rearrangement of genes e.g. RET, NTRK, ALK leading to fusions affecting receptor tyrosine kinase (RTK) protein. FvPTC: follicular variant of PTC.*

Each of the above driver mutations appear to be mutually exclusive, and are associated with individual histological pictures, gene expression and thus clinical characteristics (20). For example *BRAF*-type PTC, associated with classical or tall-cell variants histologically, has increased radio-iodine therapy resistance, lymph node metastasis and recurrence rates post-thyroidectomy (10) compared to *BRAF* wild-type PTC. In contrast to this, follicular variants of PTC that are encapsulated, associated with *RAS* variants, are felt to be of such low malignant potential they are now termed non-invasive follicular neoplasm with papillary-like nuclear features (NIFT-P) and have been reclassified as having uncertain malignant potential (22).

These different driver mutations in PTC have varying levels of impact on the MAPK pathway output (**Figure 1-4**). *BRAF* positive variants appear to have the greatest impact followed by RTK fusions, whilst *RAS* variants the least (21). This results in differing functional, and therefore clinical, impact. The *BRAF* V600E protein is not responsive to negative feedback from ERK, which results in a much higher downstream effect of the MAPK pathway. This MAPK pathway

over-activation results in the suppression, possibly through hypermethylation, of differentiated thyroid genes such as the sodium/iodide symporter (NIS) genes, e.g. *SLC5A5*. This likely accounts for the poor clinical response to radio-iodine therapy in patients with mutant *BRAF* PTC (23). By contrast, in *RAS* variants the MAPK pathway is still inhibited by ERK negative feedback and these subtypes are more responsive to radio-iodine therapy (20, 21)

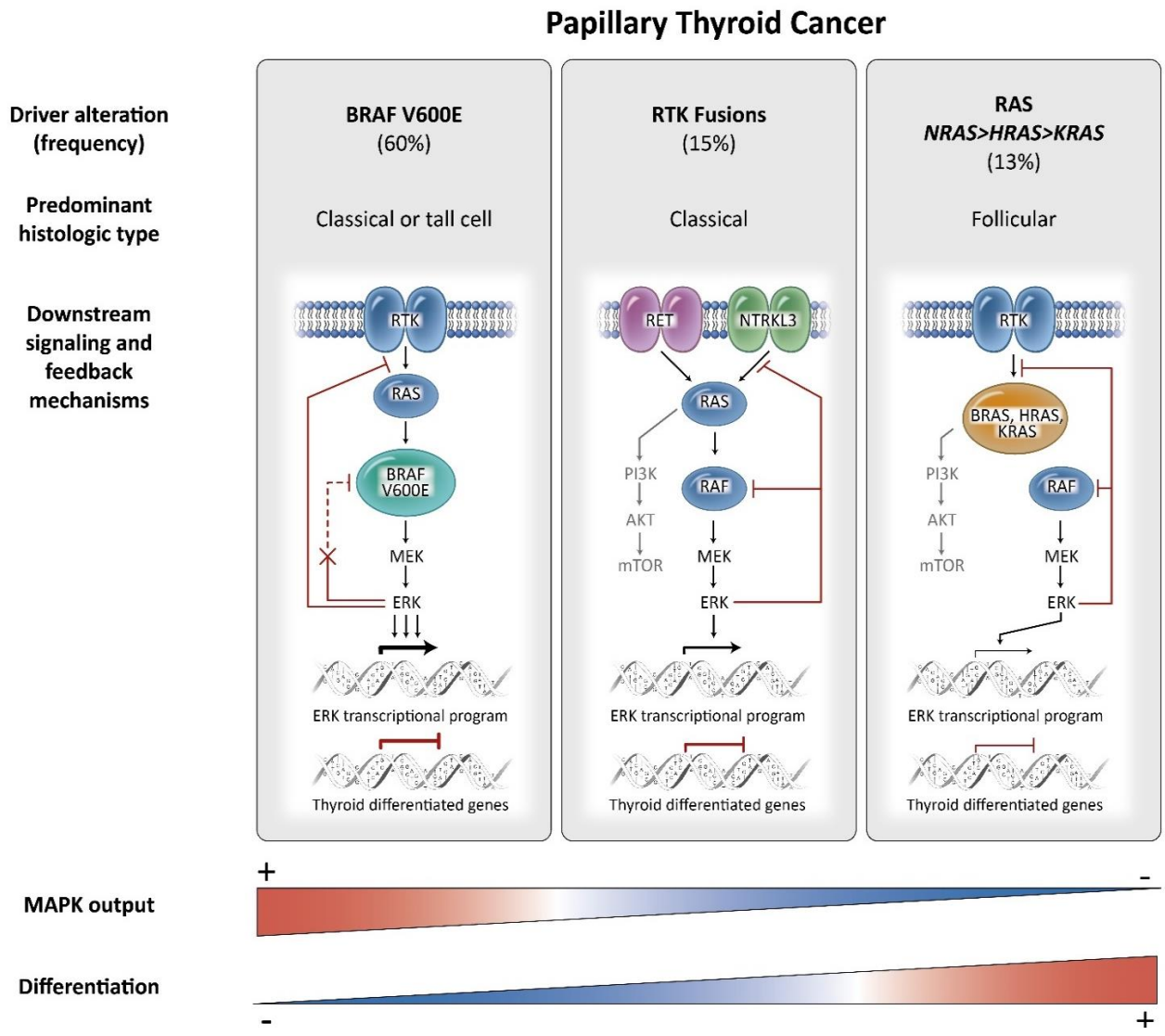

Figure 1-3. PTC subtypes defined by genomic aberration illustrated with resultant impact on MAPK pathway.
Adapted from Fagin & Wells, 2016 (21)

1.1.5.4 Follicular Thyroid Cancer

Follicular thyroid cancer appears to be a more heterogenous genetic entity than PTC. Despite this, there is still a preponderance of certain mutations. FTC appears to be associated with *RAS* variants, and are found in up to 30% of cases (24). Interestingly, a high rate of *RAS* variants (20-30%) are also found in benign follicular adenomas, suggesting that these lesions may represent precursors to FTC (25). SNV mutations involving the PI3K pathway (e.g. *PIK3CA*, *AKT1* and *PTEN*) are more prevalent in FTC. The *PAX8/PPARG* fusion oncogene is also a common event, found in up to 30% of FTCs (26). In similar fashion to PTC, where *BRAF* and *RAS* mutations appear to be mutually exclusive, in FTC *PAX8*/*PPARG* and *RAS* do not co-occur (24). It is likely that these also represent different pathological entities.

1.1.5.5 Medullary Thyroid Cancer

Sporadic MTC is strongly associated with diver mutations in *RET* and, to a lesser extent, *RAS* genes. 40-50% of sporadic MTCs harbour *RET* SNV variants, in particular p.M918T, which correlates with an increased likelihood of lymph node metastasis and poorer survival outcomes (27, 28). Whilst most cases of medullary are sporadic, up to 25% of cases are hereditary and are caused by germline *RET* mutations. Within the hereditary MTC group, a subset will be associated with MEN 2 (29). In the UK, patients with MTC are screened for germline *RET* mutations to allow for monitoring of associated endocrine abnormalities and to inform counselling of family members.

1.1.5.6 Poorly Differentiated and Anaplastic Thyroid Cancer

Both PDTC and ATC are associated with a higher mutational burden than well differentiated disease (30). It is likely that they arise from within pre-existing differentiated malignancies as a result of stepwise accumulation of mutations (31). *BRAF* and *RAS* variants are found in well differentiated cancers as well as poorly differentiated and thyroid cancers, so it has been postulated that these are likely early events in thyroid cancer progression (1). However, PDTC and ATC also often harbour additional genetic events in other loci including the *TERT* promoter region or PI3K pathway genes, as well as in genes not frequently found in differentiated cancers such as *TP53* and *CTNNB1.* Mutations involving *TP53* are found in over 50% of anaplastics (32). These mutations may be key in causing tumour de-differentiation.

1.1.6 Epidemiology of thyroid cancer

Neoplasms of the thyroid gland rank as the ninth most frequent site of malignancy. They represent the most common endocrine malignancy, and account for approximately 3.1% of all cancer diagnoses across the globe, with an estimated worldwide incidence of over 567,000 in 2018 (33). Its incidence has increased more than any other malignancy: age-standardised three year average incidence rates have risen by at least 80% for both males and females from 2005 to 2015 in the United Kingdom (34). Similar increases have been noted in other countries, including the US (35) and Australia (36). This increased incidence seems to have occurred, at least in part, as a result of better detection of small-volume PTCs due to improved imaging techniques and more frequent usage of fine needle aspiration cytology (35). This, in turn, has given rise to a substantial jump of up to four fold in observed rates of thyroidectomies (37).

Yet, despite this considerable increase in surgical intervention, mortality rates have remained relatively static over the same time frame (38) suggesting an overdiagnosis and treatment of cancers that are low-risk. The treatment and monitoring of thyroid cancer is a considerable financial burden: cost per annum in the US have been estimated at $1.6 billion, projected to rise to $3.5 billion by 2030 (39). Further, there is exposure to iatrogenic risk from the increased rates of surgery that has not had a demonstrable impact on mortality. Thus, there is a clear requirement to improve risk stratification in patients to allow better selection for surgery, radio-iodine therapy and post-treatment monitoring. Circulating tumour DNA may help in this endeavour.

1.2 Circulating DNA

1.2.1 Technological Backdrop

Recent advancements in genetic sequencing technologies have been prodigious. The initial Human Genome Project (HGP) was a public project undertaken by an international consortium that remains, to this day, the world's largest collaborative biological project. Its aim: to publish the first ever complete sequencing of the human genome. In tandem with its competitor, the privately funded Celera Genomics, it achieved its objective in 2003. Unlike many other large scale projects it was completed 2 years ahead of schedule at a lower cost than predicted, with a final estimated bill of $2.7 billion (40). This feat was mainly possible due to the rapid evolution of the sequencing techniques used in the project.

Yet, the dideoxy chain-termination sequencing method that was used by the HGP has since been supplanted by newer technologies. The more recent Next-Generation Sequencing (NGS)

technologies work in huge parallel workflows and have exponentially increased the speed of sequencing. Whole human genome sequencing (WGS) can now be down in a matter of days at a cost as low as $1,000. This rate of improvement has outpaced even Moore's Law (41). The impact of the introduction of NGS techniques is illustrated in **Figure 1-4Figure 1-4**, where a sudden dip in WGS cost can be seen circa 2007/08.

This remarkable drop in cost and time has meant that these technologies could then be viably used in the clinical setting. As such, new diagnostic entities are emerging that potentially have a powerful role in shaping clinical diagnostic capabilities.

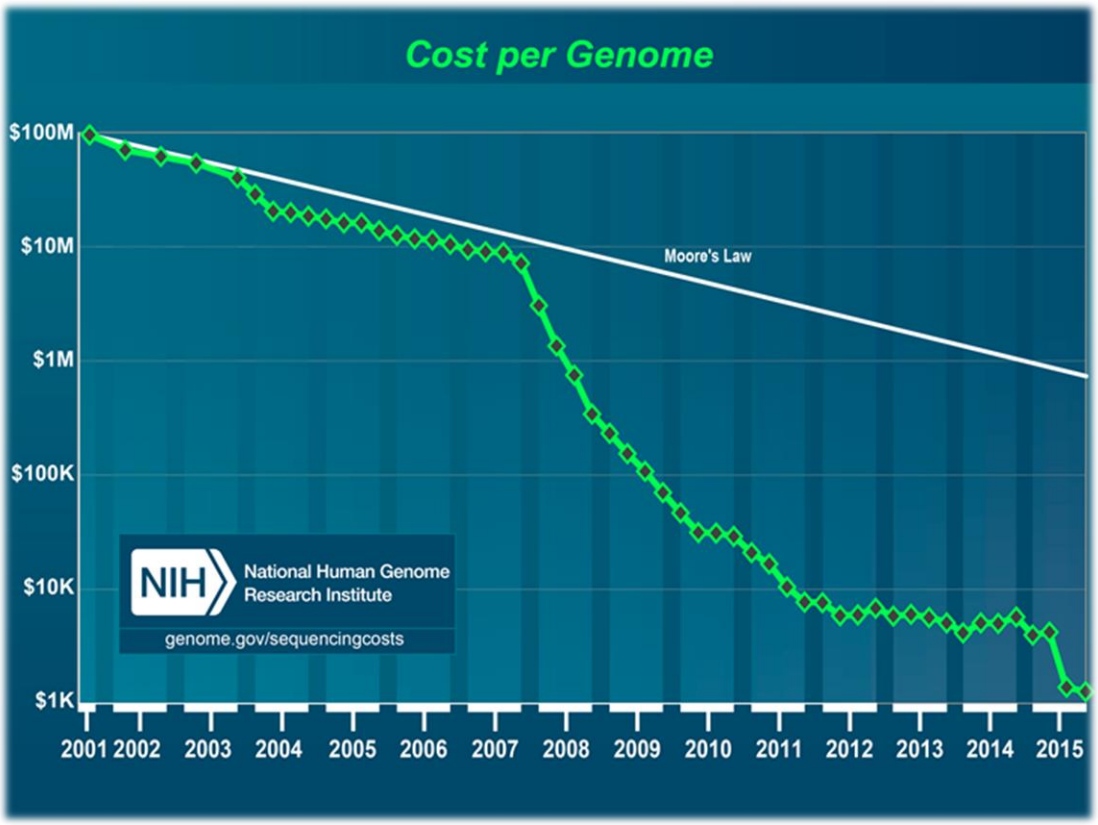

Figure 1-4. Cost of whole genome sequencing over time*. Moore's Law: the observation that transistor density in a circuit double every 2 years and costs falls in tandem. From www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost*

1.2.2 Cell-free DNA

Circulating DNA, or cell-free DNA (cfDNA), is highly fragmented genomic DNA released from cells into the circulation. The exact mechanisms through which this occurs is not yet understood, but likely originates from cells undergoing either necrosis, apoptosis or through active secretion (42-44). Cell-free DNA is detectable in plasma, but also in other body fluids, including urine (45, 46) and saliva (47). It appears to be rapidly cleared from the circulation and has a half-life of less than 3 hours (48-50).

Cell-free DNA fragments are approximately 167 base pairs in length, corresponding to the length of DNA wrapped around a nucleosome complex (51). Whilst cell-free DNA fragments released by apoptotic cells appear to be of relatively uniform size, this contrasts with necrotic cells. In necrotic cells, seen more abundantly in malignant states, cfDNA fragment size range is broader. This has been shown in a variety of solid tumours (52-54) and may be reflective of the more disordered mode of cell death occurring in necrotic cells, allowing for more random interaction between DNA and DNases.

Haematopoietic cells are likely to represent the major source of cfDNA in plasma, and levels exist at a steady state within the circulation in healthy individuals (55). However, concentration has been shown to increase in response to cellular injury that results in necrosis. This may be due to a variety of processes ranging from cancer, end-stage renal failure (56), myocardial infarction (57) and even exercise (58).

1.2.3 Historical perspective

The presence of free nucleic acid in the circulation was first demonstrated in 1948 (59), although the importance of this was not well understood at the time. It is remarkable that this discovery pre-dated even the publication of the double helix structure of DNA by Watson and Crick in 1953 (60). Further progress was not made until 3 decades later when a link between circulating DNA and cancer was made in 1977. Leon *et al* (61) found higher levels of free DNA in the serum of patients with cancer compared to those without, and in particular those with metastatic disease. In 1989, the first study was published demonstrating similarities between the DNA in the circulation to that in a primary tumour, thus linking the origin of a portion of circulating DNA to cancer (62)

The presence of cfDNA has been exploited in fields other than oncology and already has a well-established role in foetal medicine, where examination of maternal blood for cell-free foetal DNA (cffDNA) has been performed for some time. The presence of cffDNA in maternal circulation was demonstrated in 1997 (63) and has since been used for pre-natal diagnostic purposes, including early determination of foetal sex and screening for genetic abnormalities such as aneuploidies (e.g. Trisomy 21) (64).

Partly due to poor understanding of the genetic alterations involved in cancer and also the technologies used in genotyping being initially slow and expensive, exploitation of cfDNA in oncology had been relatively limited. Plasma cfDNA is present at very low concentrations and early detection methods struggled with this. However, the advent of sequencing methods, initially dideoxy chain terminating introduced by Sanger in 1977 (65), followed by NGS,

coupled with advances in other molecular assays such as digital droplet PCR (dPCR) allowing detection of much smaller quantities of DNA, allowed for the genotyping of cfDNA present in plasma. Once it was possible to correlate mutations found in the cfDNA to those found in the primary tumour, enough traction was gained to investigate the potential of circulating DNA as a biomarker in cancer.

1.2.4 Circulating tumour DNA

Tumour cells also shed fragmented genetic material into the circulation. Therefore, in patients with cancer, a proportion of circulating DNA will have originated from the tumour. Fragments of genomic DNA originating from the tumour present in the circulation are referred to as circulating tumour DNA (ctDNA). These genetic alterations are not present in non-cancerous cells and therefore ctDNA has the potential to be an exquisitely specific marker for cancer.

Analysis of ctDNA potentially enables the tracking of tumour dynamics in a more contemporaneous fashion than conventional tissue biopsies and has several potential important applications. The genetic data contained therein may offer crucial genetic information about an individual's cancer characteristics and thus ctDNA diagnostics may represent a step towards genuine personalised medicine. Furthermore, obtaining tumour tissue for DNA extraction is not always possible and may be associated with significant risk. Circulating DNA, however, is extracted from plasma obtained from simple venepuncture with almost no risk to the patient.

1.2.5 Potential applications of ctDNA

1.2.5.1 Tumour heterogeneity

Cancers may exhibit substantial genetic heterogeneity, both within the tumour and between primary and metastatic tissues (66-68). As such, traditional solid tumour biopsies may present DNA for molecular analysis that is not fully representative of a cancer's driver mutations. This may lead to inaccurate prognosis, or therapy choice that is not effective against all subclones.

In contrast, through sampling genetic material released by all malignant tissues, ctDNA may provide a more complete representation of the mutational landscape across the patient's disease burden. Due to the minimally invasive nature of obtaining ctDNA, whether that be from blood, urine or saliva, repeated biopsies are also much more feasible compared to conventional solid tissue sampling. This may allow for the emergence of new clones to be detected more easily (69) as disease progresses compared to traditional methods.

1.2.5.2 Response to therapy

1.2.5.2.1 Minimal Residual Disease

Minimal residual disease (MRD) is a well-established concept in relation to haematological malignancies. It had been defined as the presence of residual leukaemic cells undetectable by conventional morphological methods using microscopy that can be monitored using more sensitive techniques such as real-time quantitative PCR and flow cytometry (70). The term has since crossed over to solid tumours, to describe scenarios where tumour cells are not detectable through either clinical assessment or imaging but are with molecular testing.

An initial small study of 18 patients with advanced colorectal cancer (48) highlighted the potential use of ctDNA to identify patients with residual micro-metastatic disease after surgery. In this study, ctDNA was found in all patients prior to surgery with curative intent, but no recurrence was seen in patients in whom it was no longer detectable afterwards (n=4).

There has since been a drive towards establishing detection of ctDNA as a biomarker for MRD in patients following primary treatment. There is now mounting evidence to support this approach: using ctDNA as a "test of cure" in early solid, non-metastatic cancers in several tissues, including breast (71), colon (72) and lung (73). If confirmed by prospective studies, then detection - or absence - of ctDNA could allow for tailoring of adjuvant therapies to reduce relapse risk or alternatively de-escalation of treatment and monitoring if deemed low risk. In differentiated thyroid cancer, for instance, this could allow for downgrading of patients post-surgery into a category not requiring radio-iodine ablation.

1.2.5.2.2 Monitoring response to systemic therapies

Multiple biomarkers are used across a range of cancers in current clinical practice to measure treatment response or progression of disease. These include prostate specific antigen (PSA) in prostatic malignancies and cancer antigen 125 (CA 125) in ovarian. In thyroid cancers thyroglobulin (Tg) is used to monitor DTC subtypes, whereas calcitonin and carcinoembryonic antigen (CEA) are used in MTC.

These traditional thyroid cancer biomarkers, however, have several drawbacks limiting their clinical utility. These are discussed in more depth in chapter 3 but include lack of specificity, reduced accuracy in representing disease burden in the presence of biomarker antibodies and difficulty in interpretation when residual thyroid tissue is present.

In other cancers, ctDNA levels have been shown to increase as a result of disease progression and decrease in response to treatment (74-76) and it may prove to be a more specific and reliable biomarker than current protein-based markers. This would be of significant clinical benefit in thyroid cancer where markers are sub-optimal for the above reasons.

1.2.5.3 Therapeutic target identification

Multiple therapies exist that target specific cancer-associated mutations. Examples include cetuximab, an *EGFR* inhibitor used in colorectal cancer, and alectinib, an anaplastic lymphoma kinase (ALK) inhibitor used in NSCLC disease with *ALK*-fusion genes. The efficacy of these targeted therapies may be improved by selecting patients with relevant mutations. Gefitinib, for example, has better outcomes in pulmonary adenocarcinoma patients whose tumour has *EGFR* mutations (77). Molecular screening programmes aim to select, based on molecular characterisation of tumours, which targeted therapy is most appropriate in individuals (78-81).

In most studies this genotyping is performed on the primary tumour, however this may not be optimal. Limitations with this tumour-centric approach have been demonstrated and are likely due to significant genetic heterogeneity that may exist between metastatic and primary

tissue. Response rates to buparlisib, a PI3K inhibitor, in breast cancer was better predicted by genotyping of *PIK3CA* in ctDNA at trial entry than in the primary tumour (82). Consequently, the effectiveness of targeted therapies (TT) may be improved by selecting patients based on contemporaneous ctDNA analysis instead of historical primary tumour, particularly in those with advanced metastatic disease.

Targeted therapies also exist for thyroid cancer. Four have gained US Food and Drug Administration (FDA) and European Medicines Agency (EMA) approval for treatment: vandetanib and cabozantinib in MTC, sorafenib and lenvatinib in radio-iodine refractory DTC. These relatively "dirty" drugs are multi-kinase inhibitors that exert their effects through inhibition of a variety of kinases involved in the MAPK signalling cascade and in angiogenesis. More specific therapies that may be of benefit in thyroid, such as BRAF inhibitors (dabrafenib, vemurafenib) and MEK inhibitors (trametinib, selumetinib) are used in other cancers e.g. malignant melanoma (80) and lung cancer (81). However, trials are still awaited assessing their efficacy in DTC or MTC. Two trials are currently actively assessing the use of two selective RET-inhibitors: BLU-667 (ARROW trial, NCT03037385) and LOXO-292 (LIBRETTO-001 trial, NCT03157128), both of which have shown promising early results (83, 84)

As yet, few data exist on which mutations in thyroid tumours of any subtypes respond best to these targeted therapies. This is reflected in the lack of molecular screening programmes for both DTC and MTC and is something that will need to be addressed in future trials. One exception is ATC where, based on the results of a clinical trial (NCT02034110) FDA approval was granted in 2018 for a combination therapy to be used in unresectable, metastatic ATC

positive for the *BRAF* V600E variant. Patients treated with a dabrafenib in combination with trametinib, showed an overall response rate of 69% (85).

As knowledge increases of which molecular characteristics respond best to individual therapies in thyroid malignancies, molecular screening may become useful for drug selection and it is likely that ctDNA could represent a viable source of tumour DNA for this purpose.

1.2.5.4 Detection of resistance to therapy

Genomic mechanisms of resistance to therapy have been described in various cancers, and these have been detected in ctDNA. In breast cancer, oestrogen receptor gene *(ESR1)* mutations are known to confer resistance to aromatase inhibitors (86) and can be detected in the plasma (87). Likewise, most patients with non-small cell lung cancer (NSCLC) who respond to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor (TKI) therapy will develop resistance. The T790M *EGFR* variant is the most common cause of resistance to TKI therapy in NSCLC patients (88), and is also detectable in ctDNA (89, 90). Recently, Roche gained FDA approval in 2016 for the Cobas® EGFR product (91). This tests for resistance-conferring variants in the ctDNA of NSCLC patients using real-time PCR when tissue biopsy is not available, guiding treatment switch to non-resistant alternative therapy such as osimertinib.

Thus, ctDNA could be used to detect the emergence of mutations associated with resistance to a particular therapy. This may permit earlier discontinuation of a potentially toxic,

ineffective drug and switch to an effective treatment if available. Furthermore, it could do so without necessitating repeated, potentially harmful, tissue biopsies. Whilst currently not known, it is likely that resistance-conferring variants to targeted therapies in thyroid cancer will be described in the future, and their detection in plasma of emerging resistance could allow for personalised therapeutic decision making.

1.2.6 Limitations and difficulties of ctDNA

Detection of ctDNA poses considerable diagnostic challenges: it is highly fragmented, often at extremely low concentration and may represent only a tiny fraction of total circulating DNA, significantly increasing background noise. Concentration may be as low as a single copy per mL of plasma, and allelic fractions versus wild-type DNA may be under 0.01%. Recently established highly sensitive techniques, such as digital PCR, and more cutting-edge approaches such as unique molecular identifiers (UMIs) for NGS platforms, permit rapid and reliable genotyping of circulating nucleic acids that are able to cope with such low concentrations and high signal-to-noise ratio, thus allowing for meaningful investigation into the potential clinical utility of ctDNA.

Yet despite these techniques, detection rates in patients appear to peak at around 70-80%, even in advanced metastatic disease, and as low as 48% in early disease (92). This holds true across cancers originating in different tissues, including breast, lung and colorectal. Non-detection in up to a quarter of patients with the most severe disease poses a significant

obstacle to ctDNA gaining wide acceptance as a universal biomarker. It is likely, however, that the sensitivities of assays will continue to further improve and detection rates will increase.

Yet, even with improved detection, sampling error may continue to be problematic. As the concentration in plasma can often be very low, the stochastic effect of sampling relatively small amounts of plasma will become an issue if greater detection rates are desired. A typical venepuncture sample used in studies may require 20 mLs of blood, which can yield 4-8 mLs of plasma. If the concentration of ctDNA is 1 copy per mL or lower the probability of acquiring a blood sample without capturing tumour DNA is not insignificant. There is an obvious clinical ceiling to how much blood can be sampled before causing harm to the patient.

1.2.6.1 Clonal Haematopoiesis of Indeterminate Potential

Genetic alterations in haematopoietic cells are known to increase with age and these may act as a confounder of variants detected in ctDNA. Multiple recent studies have demonstrated an increasing prevalence of somatic mutations in peripheral white blood cells with age. Two large WES studies (93, 94) showed that SNVs and indels are detectable in the DNA of circulating blood cells from patients with no known haematological malignancy at the time of sampling. Whilst rare in individuals under the age of 50, this rises significantly with age: 11.7% of those aged 80-89 years (**Figure 1-5)** have detectable mutations (93).

The impact of the presence of these variants in patients with no known cancer is not yet fully understood, although detection is associated with an increased risk of haematological

malignancy and mortality. It has been postulated that these mutations are found in a sub-population of blood cells clonally derived from a haematopoietic stem cell (HSC) that has acquired the variant spontaneously. Thus, the phenomenon is described as clonal haematopoiesis of indeterminate potential (CHIP) (95).

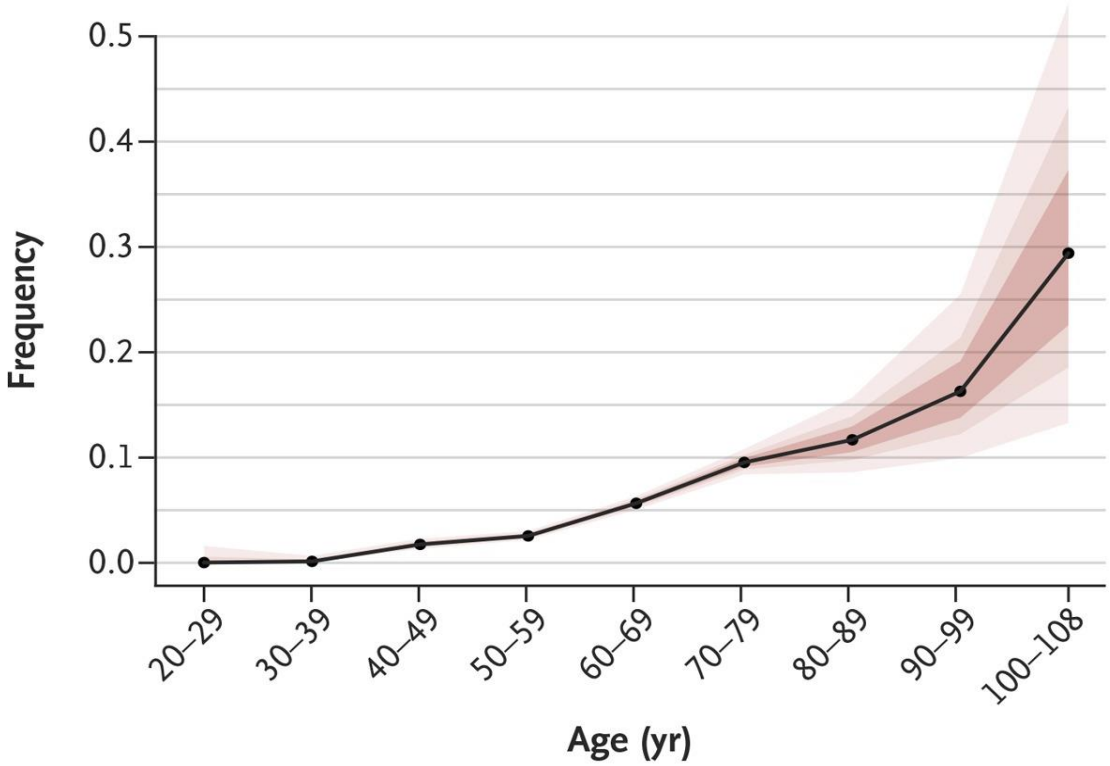

Figure 1-5. Prevalence of somatic mutations according to age. Coloured bands, in increasingly lighter shades, represent the 50th, 75th, and 95th percentiles. From: Jaiswal *et al.(93)*

In the above studies, the source DNA material for sequencing was obtained from genomic DNA that was extracted from circulating white blood cells (WBC). However, it is highly plausible that these WBCs also release cfDNA into the circulation and could contaminate plasma derived ctDNA samples with these CHIP alterations. This therefore raises questions on the inferred specificity of detected mutations in ctDNA originating from a primary tumour of interest.

Concurrent genotyping of genomic material obtained from buffy coat samples at the time of plasma sampling for ctDNA is usually performed in ctDNA studies, to act as control for germ-line mutations, and this would reduce the risk of CHIP confounding results. However, if ctDNA were to be used to track patients over lengthy periods of time, which may be a requirement in relatively indolent thyroid cancers, this may need to be done repeatedly to ensure detected mutations are not haematopoietic in origin. So far, the variants most frequently detected in CHIP are in genes often associated with haematological malignancy but not in thyroid cancer (*DNMT3A*, *TET2* and *ASXL1*), although *TP53* variants have been found. Further work is required to better understand the implications of CHIP in ctDNA.

1.2.7 Sources of circulating DNA

1.2.7.1 Alternative sources of circulating tumour DNA

1.2.7.1.1 Plasma versus Serum

Circulating DNA is present in both the serum and plasma of peripheral blood. Both are fluids in which the cellular contents have been separated through centrifugation. In plasma, however, the clotting process has been inhibited, e.g. using ethylenediaminetetraacetic acid (EDTA), whereas in serum clotting has been allowed to occur. Clotting results in the release of genomic DNA through lysis of a proportion of white blood cells. Genomic DNA release further dilutes ctDNA which increases background noise signals, thus making detection harder. As leaked, cell-free, genomic DNA has been shown to be significantly higher in serum samples, plasma has become the preferred sample fluid in the majority of studies (96, 97).

1.2.7.1.2 Trans-renal DNA

Circulating DNA is excreted via the kidneys and its presence has been demonstrated in urine (98). Since then, the use of trans-renal DNA (trDNA) has started gaining traction in some cancers, particularly lung and colorectal, where good concordance has been shown between urine and plasma derived ctDNA (45, 46, 99) . Sampling of trDNA from urine specimens obviates the need for multiple venepunctures required for plasma sampling, and hence potentially represents a truly non-invasive biopsy. The use of trDNA for tumour characterisation is still in its infancy but may prove to be a useful alternative source in the future, and in particular may be useful in overcoming the stochastic effect associated with blood sampling. Issues with contamination from epithelial cell derived DNA will also need to be overcome.

1.2.7.2 Circulating tumour cells

Tumour-derived epithelial cells are identifiable in the circulation. These circulating tumour cells (CTC) offer an alternative source of tumour genetic material that could also allow for molecular characterisation. CTCs are present in very low concentration in the blood: 1 mL of blood may contain as little as 1 CTC, but up to 10^9 red blood cells (100). In order to enrich samples from such low concentrations, CTCs must be identified and separated from normal cells. This enrichment can be achieved using a variety of techniques based either on cell-surface labels, such as immuno-magnetic separation (e.g. CellSearch) or biophysical properties, such as Isolation by Size of Epithelial Tumour cells (ISET) or newer micro-fluidic techniques (101, 102). Cytometric or nucleic acid-based techniques are then used for analysis of cells. These current techniques, however, offer relatively low yields of CTCs and therefore

ctDNA continues to be the dominant source of peripherally obtained tumour material in published studies.

1.3 Theme of thesis

The sections above set out the background to this thesis, which focuses on circulating tumour DNA and its potential role as a biomarker in thyroid cancer. Published data on ctDNA in thyroid cancer are relatively sparse compared to several other solid tumours, discussed further in chapter 3, and the role that it may play clinically is still uncertain. For ctDNA to gain momentum as a clinically useful biomarker in thyroid cancer it needs to demonstrate significant detectability, which has been lacking in studies thus far. It is within this context that the experiment in the following chapter was conducted: through a multi-mutational approach and making use of more sensitive assays improved detection rates were hoped for. This was built on in the subsequent experiments where I developed a validated targeted gene panel and conducted a pilot study on mechanisms of resistance as proof-of-principle to establish an experimental workflow. These will both be useful for future studies.

1.4 Hypotheses and aims of thesis

1.4.1 Hypotheses

1. That ctDNA is detectable in the plasma of patients with advanced thyroid cancer and may represent a novel biomarker for monitoring disease burden
2. That a thyroid specific targeted gene panel would offer superiority over a currently used generic panel (with regards detection rates, costs and time)
3. That resistance to targeted therapy in thyroid cancer is due to a genetic variant that is detectable in ctDNA

1.4.2 Aims

1.1 To conduct a multi-mutational analysis in ctDNA across a range of thyroid cancer subtypes over multiple time points

1.2 To correlate changes in ctDNA concentration with currently used biomarkers (Tg, Calcitonin, CEA) and CT imaging, with the goal of generating evidence that ctDNA detects disease progression either earlier or more reliably than current biomarkers.

2.1. To design a comprehensive targeted gene panel specific for thyroid cancer hotspots

2.2 To validate the novel panel on thyroid tumour tissue

2.3 To explore utility of this novel panel for sequencing of plasma ctDNA

3.1 To perform WES of tumour and plasma samples of patients with known acquired resistance to targeted therapies to identify potential resistance-conferring mutations

Chapter 2 Materials and Methods

2.1 Identification of Molecular Biomarkers for Thyroid cancer

2.1.1 Trial Recruitment

Identification of Molecular Biomarkers for Thyroid Cancer was a prospective observational study (CCR4343, REC reference 15/WS/0148) that was conducted at the Royal Marsden Hospital and The Institute of Cancer Research. The study recruited patients with thyroid cancer from December 2015 until July 2016. Patients targeted were adults with thyroid malignancy of any histology subtype with initial Union for International Cancer Control (UICC) Tumour Node Metastasis (TNM) clinical staging of T3-4 N0-1 M0-1. Patients provided written consent to having archival tumour tissue retrieved and analysed for the study. They also consented to regular blood draws for ctDNA analysis. Entry into the study was on a rolling basis according to patient recruitment, and not according to pre-defined clinical time points. The full trial protocol is shown in ***Appendix A***

2.1.2 Tumour DNA extraction

Archival formalin-fixed, paraffin-embedded (FFPE) tissue blocks were used for extraction where possible. If FFPE blocks were not available, then haematoxylin and eosin-stained (H&E) slides were used as an alternative. From the FFPE blocks six slides were cut by microtome: one H&E slide was reviewed by a specialist pathologist to identify the tumour and comment on cellularity and tumour content, and five 10 μM non-stained slides were used for DNA extraction. The non-stained slides were de-waxed using established protocols. The H&E slide was used as a guide for manual macro-dissection of tissue from the non-stained slides using

a single-use surgical blade. DNA was extracted using a spin column-based nucleic acid purification kit (Qiagen QIAamp DNA FFPE kit, Qiagen, Hilden, Germany). DNA clean-up was performed on samples with a concentration of <5 ng/µL using a DNA clean and concentrator kit (Zymo Research, CA, USA).

2.1.3 Plasma DNA extraction

20 mL of peripheral blood was collected from patients at 3 monthly intervals, alongside their routine blood sampling. Streck Cell-free DNA Blood Collection Tubes® (La Vista, NA, USA) were chosen over traditional K_3-EDTA tubes as they contain a cell preservative which reduces cell lysis. This allows for longer sample processing times with reduced risk of genomic DNA (gDNA) contamination from leucocytes (103, 104). Heparinised tubes were not used as it may inhibit downstream PCR processes. Time to processing of the samples was kept as short as possible to further minimise the risk of this, aiming for within 2 hours of venepuncture following established laboratory operating procedures (SOP).

On receipt in the laboratory, the blood samples were centrifuged twice for 10 minutes at 1600g. Double spinning of plasma reduces contamination from genomic DNA (105). The plasma was then aliquoted and stored at -80°C. Increasing numbers of freeze-thaw cycles can affect ctDNA (97), therefore samples were only thawed once, when downstream analysis was to be performed.

Cell-free DNA was extracted from approximately 4 mL of plasma using the QIAsymphony DSP Circulating DNA Kit (Qiagen, Hilden, Germany) on a QIAsymphony SP platform.

2.1.4 Buffy coat samples

All patients had buffy coats extracted from spun plasma sample. DNA was extracted from resuspended blood cells using the QIAsymphony DSP Midi kit and quantified using the Qubit dsDNA BR Assay kit. These were sequenced concurrently with the tumour samples to permit filtering of germline variants

2.1.5 Library preparation

Library preparation was performed using the KAPA HyperPlus kit (Roche, Switzerland) following standard operating procedures based on the manufacturer's protocol. DNA sample concentrations were quantified using a Qubit Fluorometer (ThermoFisher, Massachusetts, USA). The input amount of DNA per sample was 200 ng if the extracted sample concentration was >5 ng/µL with an average peak fragment size of >1000 base pair (bp) length. If the average peak fragment size was <1000 bp, then 400 ng was used to account for the higher fragmentation. When extracted DNA concentration did not allow for the approach described above, then a low input (50 ng) protocol was used.

Samples underwent enzymatic fragmentation using KAPA 5X fragmentation enzyme for 30 minutes on a thermal cycler at 37 degrees Celsius to achieve fragment sizes of 150-200 bp. End repair, A-tailing and adaptor ligation was then performed for each sample. Dual size selection and short-cycle amplification clean-up were performed using AMPure XP clean-up beads (Beckman Coulter, Brea, CA, USA). Pre-hybridisation library amplification was performed using 6 cycles of PCR for the 200 and 400 ng samples or 10 cycles for the 50 ng protocol under the following conditions:

PCR Program			
Step 1	98°C	45 seconds	
Step 2	98°C	15 seconds	**6 or 10 cycles**
Step 3	60°C	30 seconds	
Step 4	72°C	30 seconds	
Step 5	72°C	1 minute	
Step 6	4°C	∞	

2.1.6 Panel hybridisation

A capture-based target enrichment gene panel was used. The panel used was The Royal Marsden Hospital Gastro-intestinal panel, version 2 (RMH GI2) that targets known cancer mutation hotspots present in a wide array of solid tumours (**Table 2-1**). Although this panel was originally designed for colorectal tumours, it was chosen for this project as no thyroid specific panel existed and there was significant overlap of relevant loci across both colorectal and thyroid cancers. The panel was designed in-house and has previously been validated for clinical molecular tumour board reporting.

GI2 Panel Gene Hotspot List			
AKT1	*DOCK2*	*KIT*	*PTEN*
APC	*EGFR*	*KRAS*	*RET*
ARID1A	*ELMO1*	*MAP2K1*	*ROS1*
ATM	*ERBB2*	*MAP2K2*	*SMAD4*
BRAF	*ERBB4*	*NOTCH 1*	*TCF7L2*
ARID1A	*FBXW7*	*NOTCH2*	*TP53*
ATM	*HRAS*	*NOTCH3*	*UGT1A1*
CDKN2A	*IDH1*	*NRAS*	*VHL*
CDKN2B	*IDH2*	*PDGFRA*	
CTNNB1	*JAK3*	*PIK3CA*	

Table 2-1. *List of genes in which GI2 panel covered known cancer hotspots*

Roche SeqCap® kits were used for panel hybridisation. The GI2 panel consists of single-stranded biotinylated RNA baits that were hybridised to each sample for at least 16 hours at 47 degrees Celsius on a thermal cycler. Capture of baits was then performed with kit-supplied magnetic beads containing streptavidin, allowing binding through streptavidin-biotin bonding. Following capture samples underwent post-hybridisation PCR amplification for 11 cycles:

PCR Program			
Step 1	98°C	45 seconds	
Step 2	98°C	15 seconds	**11 cycles**
Step 3	60°C	30 seconds	
Step 4	72°C	30 seconds	
Step 5	72°C	1 minute	
Step 6	4°C	∞	

Samples were then cleaned with AMPure XP beads. Quality control and quantification of samples were performed using the KAPA Library quantification kit on a TaqMan quantitative PCR platform according to manufacturer specifications.

2.1.7 Sequencing and Bioinformatics

The resulting indexed libraries, together with positive (Promega DNA) and negative (no template) controls, were pooled and sequenced according to manufacturer user-guides on either a NextSeq500 or MiSeq platform (Illumina, San Diego, CA, USA) with 75 base paired-end reading to a mean depth of >300x.

Samples were demultiplexed using Bcl2fastq2 v.2.19 (Illumina) and aligned to the Genome Reference Consortium Human Build 37 (GRCh37) using BWA-mem v.0.7.12 software. Data were filtered through an in-house pipeline that included GATK *Mutect2* version 3.5.0 (Broad Institute, Cambridge, MA, USA) for single nucleotide variant (SNV) and insertion/deletion (Indel) identification. Oncotator v.1.9.9.0 was used for annotation of resulting variant call files (.vcf). Manta v.0.29.6 was used for identification of structural variants. Additional in-house filters applied to variant calling included at least 10x depth for variant calls and a custom Gnomad filter to remove common SNVs. Deletion and amplification thresholds of <0.5x and >2.4x fold change, respectively were used for copy number calling.

2.1.8 Plasma genotyping

Plasma samples were genotyped using the QX-200 droplet digital PCR system (Bio-Rad, California, USA). Custom Taqman assays were used (Bio-Rad & Applied Biosystems, Thermo Scientific) according to the variants discovered from tumour tissue sequencing. For each plasma time-point, 10.45 µL of DNA was placed per well in at least 2 wells and interrogated with the custom assays. Mean cfDNA concentration of samples was 0.43 ng/µL (range 0.079 – 1.47). For each patient-specific assay one No Template Control (NTC), one wild-type control (fragmented Promega DNA at 1 ng/µL) and one positive control (patient's own tumour DNA) were used. cfDNA samples were emulsified on a droplet generator. Emulsified PCR reactions were run on 96-well plates on a C1000 Touch Thermal Cycler (Bio-Rad), using the following conditions:

PCR Program			
Step 1	95°C	10 minutes	
Step 2	94°C	30 seconds	**40 cycles**
Step 3	60°C	60 seconds	
Step 4	98°C	10 minutes	
Step 5	4°C	∞	

The temperature ramp increment was 2.0°C/s for all steps. Samples were read on a QX-200 ddPCR droplet reader as per manufacturer's guidelines. Results were analysed using Bio-Rad's QuantaSoft software v.1.7.4, and individual cut-offs were established for each assay. A

minimum of 20,000 droplets across both wells was required for a valid test. A minimum of 2 single FAM positive droplets with a number of double positive droplets lower than single FAM droplets was required to make a positive call. If no positive result was obtained in 2 wells, then the assay was repeated for a further 2 wells.

2.1.9 Conventional biomarker assays

The Access Thyroglobulin and Thyroglobulin Antibody II chemiluminescent immunoassays (Beckman Coulter, Brea, CA, USA) were used for Tg and anti-Tg antibody quantification, respectively. The calcitonin assay was done with the LIAISON® Calcitonin II-Gen one-step sandwich chemiluminescence immunoassay (DiaSorin, Saluggia, Italy). Quantitative CEA assaying was performed using the ARCHITECT chemiluminescent microparticle immunoassay (Abbott Laboratories, IL, USA).

2.1.10 Imaging

Contrast-enhanced CT scanning was used for axial imaging in this study, with views obtained of the head, neck, thorax, abdomen (liver imaging was initially triple phase at baseline, dual phase thereafter for patients with MTC) and pelvis. Progression was defined as per RECIST criteria (106)

2.1.11 Statistics

Prism 8 software (GraphPad, San Diego, CA, USA) was used for simple statistical testing. More complex modelling was performed by Dr Jennifer Liu, senior research statistician, using R (R Core Team, 2014). Statistical tests used were as follows:

Section 8.4.3: Association testing between mutations and disease subtype was performed with Chi-squared testing for overall correlation, and then generalised linear models for individual mutations with disease subtype as the outcome and mutations as the predictor.

Section 8.4.5*:* Testing of correlation between variant allele fraction detected by NGS and dPCR was performed using Spearman rank correlation test.

Section 8.4.6: Determining if ctDNA detection is higher in patients with metastatic disease was tested using a generalized linear model with binomial outcome and disease status as a factor, adjusting for age and sex. Mann-Whitney testing was performed on the median concentration of ctDNA concentrations in patients with metastatic disease compared to non-metastatic (local recurrence or no structural disease) patients.

Section 8.4.7: Testing of correlation between ctDNA and conventional marker concentration was performed using logged concentration values and Spearman rank correlation test.

Section 8.4.8: The Kruskal-Wallis test was used for comparison of median ctDNA concentrations to disease progression (PD vs SD & PR). Determining the effect of changes in ctDNA concentration on disease progression (PD) was modelled using a

linear mixed model using the following parameters: normalised ctDNA concentration, time and interaction between time and ctDNA concentration.

2.2 Development and validation of thyroid specific gene panel

2.2.1 Literature review

Published literature was reviewed to ascertain common cancer hotspots relevant to thyroid cancer that would be desirable to include in a novel targeted panel. The PUBMED online database was used to identify relevant articles, and this search was performed from April to May 2017. Suitable papers were used to identify the loci of genomic aberrations that would be included in the panel design. Once these were identified, genomic coordinates were chosen that framed these loci of interest.

2.2.2 Design of panel

Once the loci of interest were decided, the panel was designed using the ThermoFisher Scientific Ion AmpliSeq Designer tool and submitted for production. The panel was optimised for shorter DNA fragments, to allow for sequencing direct from plasma. It was supplied in 2 pools of amplicons to avoid overlapping and cross-reactivity of amplicons. The total coverage size was 20.47 kb. The final panel had a 92.36% coverage of the submitted loci. The achieved coverage in the panel was manually curated prior to purchase to ensure optimal coverage for crucial hotspots. A summary of the covered loci is shown in **Table 2-2**.

ThyMa Panel Gene Hotspot List			
Gene	**Exons**	***Gene***	**Exons**
BRAF	15, 16	*AXIN1*	2-5, 7-11
RET	2, 10, 11, 15, 16	*GNAS*	8, 9
TSHR	8, 10, 11	*APC*	6-16
TERT	Promoter	*KRAS*	2-5
NRAS	2-4	*PTEN*	4-8
TP53	2-10	*AKT1*	2, 4
KMT2C	7-18	*EIF1AX*	1-7
HRAS	2, 3	*CHEK2*	4, 12, 16
PIK3CA	2, 5, 10, 21	*ALK*	7, 8, 12, 16, 19, 23, 25
CTNNB1	3, 11, 15	*IDH1*	6-4

Table 2-2. Summary of genes and exons covered by the novel panel

2.2.3 Validation experiment

Once obtained, the novel ThyMa panel was submitted through a validation experiment. The primary aim of this experiment was to ensure non-inferiority to the RMH GI2 panel used in the ctDNA project and also to explore its functionality in sequencing from plasma.

2.2.3.1 Samples used for validation

All the tumours used in the Chapter 3 study were put forward for use in the validation experiment.

Plasma samples were eligible if ctDNA was detected in the previous experiment (Chapter 3) and had sufficient DNA remaining. Samples with a variety of fractional abundances (1-26.69%) were selected. Where possible, samples taken from different time points in the same patient were used to enable comparison between the panel and ddPCR at tracking ctDNA titres.

DNA extracted from two cell lines was also used. TPC-1 and HTH-74 are human thyroid cancer cell lines, the former from papillary carcinoma and the latter from anaplastic carcinoma. Both are known to harbour the *TERT* variant c.1-124C>T in the promoter region and were used as positive controls for this variant. As this variant is a common mutation in thyroid cancers and is located in a region that is difficult to design primers for, due to multiple homopolymers and high GC content (78%), it was felt important to include known positives to be run with the new panel.

2.2.4 Library preparation

Ion Torrent-compatible libraries were constructed using the Ion Ampliseq Library Kit v2.0. Initial multiplex PCR enrichment was performed in two wells per sample, as this was a two-pool panel. This was performed with 2 µL of 5X Ion Ampliseq HiFi Mix, 5 µL of 2X ThyMa Ion Ampliseq Primer Mix made up to 10 µL with DNA and nuclease-free water. Input amount of DNA was 10 ng for tumour samples, and 3 ng for plasma. Cycling conditions for the multiplex PCR were dictated by the Ion Ampliseq Library Kit v2.0 and varied according to the number of primer pairs and input quantities of DNA. Primer digestion was subsequently performed with post PCR volumes having 1 µL FuPa reagent added and being subjected to 50^{o}C for 10 minutes, 55 ^{o}C for 10 minutes then 60 ^{o}C for 20 minutes. Indexed adapters were ligated by adding 1 µL Ion Ampliseq Adapters, 2µL Switch Solution and 1 µL DNA Ligase followed by 22 ^{o}C for 30 minutes and 72 ^{o}C for 10 minutes in a thermal cycler. The resultant library was purified using 1.5X AMPure XP beads and subjected to two washes with 70% ethanol prior to quantification.

2.2.5 Library quantification

Duplicate reactions of 1:400 library dilutions were tested to estimate library molarity for pooling. Reactions consisted of 10µl Ion Torrent Ion Library TaqMan qPCR Mix, 10µl Ion Library TaqMan Quantitation Assay and 9µl of library dilution. Standards of 6.8pM, 0.68pM and 0.068pM of *E. coli* DH10B Control Library and NTC wells were included in triplicate. The following program was run on a QuantStudio 6 Flex from ThermoFisher:

50°C	2 minutes	Hold
95°C	20 seconds	Hold
95°C	1 second	40 x cycles
60°C	20 seconds	

Libraries were finally pooled at 100pM for loading on the Ion Proton chip.

2.2.6 Sequencing & Bioinformatics

The pooled samples were loaded onto a Proton P1 chip and sequenced on an Ion Torrent Proton platform (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing reads were aligned and BAM files generated for the IonTorrent libraries using the Torrent Suite software. Variants were called using a combination of MuTect2 (107), VarDict (108) and the torrent caller. Samples were annotated using ANNOVAR (109) followed by manual curation of the detected variants.

2.2.7 dPCR validation of novel variants detected in tumour

Where possible, new variants identified through sequencing in tumour samples were confirmed using ddPCR, as described in **section 2.1.8.**

2.2.8 Plasma sequencing sample selection and analysis

An ability to sequence direct from plasma would be advantageous for future studies, therefore the ThyMa panel was also run with plasma samples. However, this is technically more challenging due to significantly lower concentration and higher fragmentation of cfDNA. Two experimental approaches were used, to assess sequencing from plasma using the ThyMa panel in light of these challenges, and then to explore potential use of direct plasma sequencing in detection of new developing sub-clonal mutations:

1. *Limits of detection*

   Sequencing data were reviewed to establish if variants detected using dPCR were seen. If detected, then variant allele frequency was compared across both methods to assess for correlation. The detection of variants according to concentration of DNA in the plasma and tumour allele frequencies were observed to establish crude limits of detection.

2. *Novel variants*

   To identify novel variants in plasma samples that were not detected in the original tumour samples, plasma sequencing data was compared to the sequencing data from the tumours obtained in the chapter 3 study. Variants present only in the plasma samples were treated as potentially representing new tumour mutations that had

arisen at a later stage in the disease process. These candidates were then manually filtered to include only exonic, non-synonymous variants. A read depth of at least 100x and a variant AF of 2% or above was required for a positive call. Variants with allele fractions of 45-55% were excluded as, in the absence of confirmatory buffy coat sequencing, these were likely to represent germ-line single nucleotide polymorphisms (SNPs)

2.3 Mechanisms of resistance in advanced thyroid cancer

2.3.1 Samples

Clinical records from the patients enrolled in the Chapter 3 study were reviewed to identify any patients on targeted therapy (TT) that had plasma samples taken *after* documented disease progression. Disease progression after TT was taken to indicate resistance to the TT. This time point was chosen as it was felt that if resistance was due to a novel mutation, it would be present in the ctDNA of this sample. Matched plasma samples, pre- and post-TT were used when possible for the shorter interval between samples. When a pre-therapy plasma sample was not available then a tumour FFPE sample was used instead as a DNA source.

2.3.2 Library Preparation

FFPE tumour samples underwent mechanical fragmentation using an E220 Focused-ultasonicator (Covaris, Chicago, Il, USA) for 200 seconds, according to in-house standard operating procedures. This step was omitted in plasma samples due to the inherent short fragment size of circulating DNA.

Libraries were then constructed using the *Agilent SureSelect XT* kit (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's protocol. Firstly, end-repair, A-tailing and adaptor ligation was performed. Samples then underwent a pre-hybridisation PCR amplification of 10 cycles. One sample had <15 ng DNA post-adaptor ligation and so underwent a further 2 cycles. The PCR cycle conditions used were as follows:

PCR Program			
Step 1	98°C	2 minutes	
Step 2	98°C	30 seconds	**10 cycles**
Step 3	65°C	30 seconds	
Step 4	72°C	1 minute	
Step 5	72°C	10 minutes	
Step 6	4°C	∞	

Whole exome capture was then performed using the *Agilent Human All Exon V6* baits. These baits were hybridised to samples and captured using magnetic streptavidin beads. Samples then underwent indexing and post-hybridisation PCR amplification for 7 cycles, under the following conditions:

PCR Program			
Step 1	98°C	2 minutes	
Step 2	98°C	30 seconds	**7 cycles**
Step 3	57°C	30 seconds	
Step 4	72°C	1 minute	
Step 5	72°C	10 minutes	
Step 6	4°C	∞	

After sample purification with AMPure XP beads, final quality control was performed using the TapeStation High Sensitivity D1000 kit (Agilent) to assess sample fragment size and concentration. The StepOnePlus Real-time PCR system (Thermofisher) was used to accurately quantify library concentration. Indexed libraries were then pooled together in equimolar concentrations

2.3.3 Sequencing and Bioinformatics

The pooled samples underwent paired-end sequencing on an Illumina HiSeq 2500 platform at 100 cycles for each end. The target coverage was 100x for each sample. Reads were aligned to the hg19 reference genome using BWA v0.7.15 and duplicates removed with *MarkDuplicates* in gatk4 (version 4.0.5.1) according to GATK best practices (110). Copy number was assessed using *cnvkit* version 0.9.3 (111) with allele frequencies extracted using the GATK haplotype caller. Variant calling was performed with *MuTECT2* and was additionally run in tumour only mode. Variants were annotated using *ANNOVAR*.

2.3.4 Manual curation and review of calls

Variant calls were selected for review when present across multiple patients. Resultant calls in loci where there was a low number of total reads (under 50x) were excluded. For all remaining variants published literature was reviewed to assess the currently understood function of the relevant gene. This was done to establish if there was biological plausibility that a genomic change in that gene could represent a potential mechanism for resistance to targeted therapy.

Chapter 3. Detection of ctDNA in advanced thyroid cancer

3.1 Introduction

Serum thyroglobulin (Tg) concentration is currently used as a biomarker in differentiated thyroid cancer (DTC) for detection of residual thyroid tissue, persistent disease and relapse. As mentioned previously, however, it has several drawbacks that limit its clinical utility:

- Detection of Tg in plasma is tissue-specific, not cancer-specific. It is secreted by both benign and malignant tissue, thus reducing its specificity.
- Tg levels do not accurately represent tumour burden in circumstances when Tg auto-antibodies are present, and this affects up to 25% of patients (112, 113). Similarly, in de-differentiated carcinoma Tg is no longer secreted by tumour cells and is therefore a less reliable biomarker.
- Tg levels are more difficult to interpret in patients treated with thyroid lobectomy only or who have not undergone radio-iodide ablation, as persistent thyroid tissue will continue to secrete Tg. An increasing number of patients are undergoing lobectomy only as definitive treatment for small, low risk cancers. The ongoing IoN study (NCT01398085) is assessing the need for radio-iodide ablation in lower risk DTC. It is likely, therefore that more patients will fall into this category in the future.

In medullary thyroid cancer (MTC), Tg is a redundant marker as it is not secreted by the parafollicular cells that give rise to this subtype of cancer. Thus, calcitonin and carcinoembryonic antigen (CEA) are used as biomarkers. These are also not cancer-specific and can be slow to reflect changes in disease status.

Serial imaging, using techniques such as computerised tomography (CT) or ultrasound (US) scanning, is likewise used for monitoring and detection of disease progression. However, small volume disease changes may not be readily visible on anatomical or functional imaging, which can lead to a significant lag time in detecting progression. CT scanning, a frequently used modality, leads to repeated exposure of patients to ionising radiation which, in patients with a relatively indolent disease, is an important consideration. There is also an associated financial cost to the health service of serial imaging of a large nationwide cohort of patients.

Given these limitations, new biomarkers are required to assist in the clinical management of thyroid cancer. A biomarker that could detect progression earlier with a greater specificity would be extremely useful. One that better identifies aggressive subtypes, detects recurrence at an earlier time-point and predicts response to therapies would be especially desirable.

As discussed in the introductory chapter, technological advances in genomics, such as NGS and dPCR assays, allow detection of ctDNA at extremely low concentrations within clinically appropriate time frames and cost constraints. These advances paved the way for circulating tumour DNA (ctDNA) to emerge as a viable cancer biomarker. There is now substantial evidence of clinical utility in a variety of cancers (71, 114, 115).

To date, multiple studies have looked at detection of ctDNA in thyroid cancers. However, they are heterogenous in their objectives, inclusion criteria, reported outcome measures and their methods of ctDNA detection, making collective comparisons difficult. These are summarised in **Table 3-1**. An initial study in 2010 by Chuang *et al.* (116) detected *BRAF* c.1799T>A in the

serum of 3 of 5 patients with papillary tumours that harboured the same variant. Whilst these results were clearly interesting, the size of the study was limited and subsequent studies have attempted to build on this.

Author	Year	Hist	Source	Assay	Variants	Time
Chuang *et al.*	2009	PTC	Serum	RT-PCR	BRAF V600E	1
Cradic *et al.*	2010	PTC	Plasma	RT-PCR	BRAF V600E	1
Kwak *et al.*	2012	PTC	Serum	RT-PCR	BRAF V600E	1
Pupilli *et al.*	2013	PTC	Plasma	RT-PCR	BRAF V600E	2
Kim *et al.*	2015	PTC	Plasma	RT-PCR	BRAF V600E	1
Lubitz *et al.*	2016	PTC	PBL	RT-PCR	BRAF V600E	1
Lubitz *et al.*	2018	PTC	PBL	RT-PCR	BRAF V600E	2
Condello *et al.*	2018	PTC	Plasma	RT-PCR	BRAF V600E	1
Almubarak *et al.*	2020	PTC	Plasma	dPCR	BRAF V600E	1
Zane *et al.*	2013	Mixed	Plasma	RT-PCR	BRAF V600E	1
Cote *et al.*	2017	MTC	Plasma	dPCR	RET M918T	1
Sandulache e*t al.*	2017	ATC	Plasma	NGS	Multi	1

Table 3-1. Summary characteristics of published studies of ctDNA in thyroid cancer*. Hist: histology. Source: source of ctDNA. PBL: peripheral blood lymphocyte. Assay: ctDNA detection assay used. RT-PCR: real time PCR. dPCR: digital PCR. Variants: variant detected with assay. Time: number of time points blood sampled during study.*

Thus far, studies in PTC have focused largely on detection of *BRAF* c.1799T>A in ctDNA and the results appear to be somewhat contradictory. Cradic *et al.* (117) looked at a much larger cohort (n=173) of PTC patients with a broad range of disease status at the time of study. They found that circulating *BRAF* c.1799T>A was detectable in the plasma of 8 of 42 (19%) *BRAF* positive tumours. Four other studies have demonstrated *BRAF* detection in blood (118-121), two of which demonstrated significant reductions in ctDNA titres after surgery, supporting the view that ctDNA could be a marker for disease burden (118, 120). A more recent study that also detected *BRAF* c.1799T>A in ctDNA demonstrated that levels are reflective of

tumour burden, although it is not clear from the data in what proportion of patients it was detectable (122). In contrast to these, however, three studies published by other groups were unable to detect this variant in the ctDNA of tumour positive patients (123-125).

Mixed results have been obtained in studies looking at other thyroid cancer subtypes. In 50 patients with a proven *RET* p.918M>T variant in the primary tumour, Cote *et al.* (126) detected this variant in the plasma of 32% of patients with MTC. They did, however, demonstrate a correlation between *RET* ctDNA detection and reduced survival.

In anaplastic tumours Sandulache *et al.* (127) utilised targeted gene panels to sequence both tumour and plasma on an NGS platform, instead of using RT-PCR. They found 100% detection of *BRAF* variants in ctDNA of patients with distant or locoregional disease prior to treatment and high rates for other variants in the same group (n=12). Concordance between tumour and plasma sequencing results for patients who were either post-treatment (n=7) or with no evidence of active disease (n=4) was much lower: 0.06 and 0.00 respectively. The lower concordance in the latter two groups may partly be due to the patients being disease free, but the disease-related deaths of a number of patients in these groups indicate that lack of detection does not represent absence of disease, and may be due to lack of assay sensitivity. Newer techniques that aid in accurate sequencing at the lower concentrations seen in plasma, such as unique molecular identifiers (UMI), were not used in this study.

The ctDNA detection methods used in the above studies are unfortunately not consistent, and this may partially explain the heterogenous results obtained. Differing sources of ctDNA were used and included sera, plasma, and in some studies retro-transcribed RNA from lymphocytes was used as a proxy. Furthermore, many of the studies relied on quantitative real-time PCR detection methods, which arguably have lower detection sensitivities than more recent digital PCR techniques.

Collectively, these studies either only looked at single variants or single time points. No published study has tracked multiple variants over multiple timepoints in thyroid cancer, and it is within this context that this experiment was conducted. Plasma samples were collected at multiple time points in the hope that, if ctDNA levels are representative of tumour burden, this would be reflected through observed changes in concentrations over time. To increase the chance of having a targetable variant in the ctDNA a multi-mutational approach was used, thus tumours were genotyped through sequencing with a targeted gene panel. Digital PCR, rather than RT-PCR, was used for ctDNA detection in this study as this technique was felt to be the more sensitive assay.

3.2 Hypothesis

Cell-free tumour DNA is detectable in the plasma of patients with advanced thyroid cancer and may represent a novel biomarker for monitoring of disease burden.

3.3 Aims

1. To conduct a multi-mutational analysis in ctDNA across a range of thyroid cancer subtypes over multiple plasma time points.
2. To correlate changes in ctDNA concentration with currently used biomarkers (Tg, Calcitonin, CEA) and disease progression detected on axial imaging.
3. To assess if ctDNA detected disease progression either earlier or more reliably than current biomarkers, to form the basis for a larger, multi-centre trial

3.4 Results

3.4.1 Study summary and recruited patient characteristics

52 patients were recruited to the study (**Figure 3-1**). Tissue was not available for analysis in one patient. The histological subtypes of the final 51 patients were as follows: papillary (*n* = 17), follicular (15), medullary (15), poorly differentiated (3) and anaplastic (1). The tissue used for sequencing was primary tumour in 43 cases (84%) and metastatic tissue in 8 cases (16%).

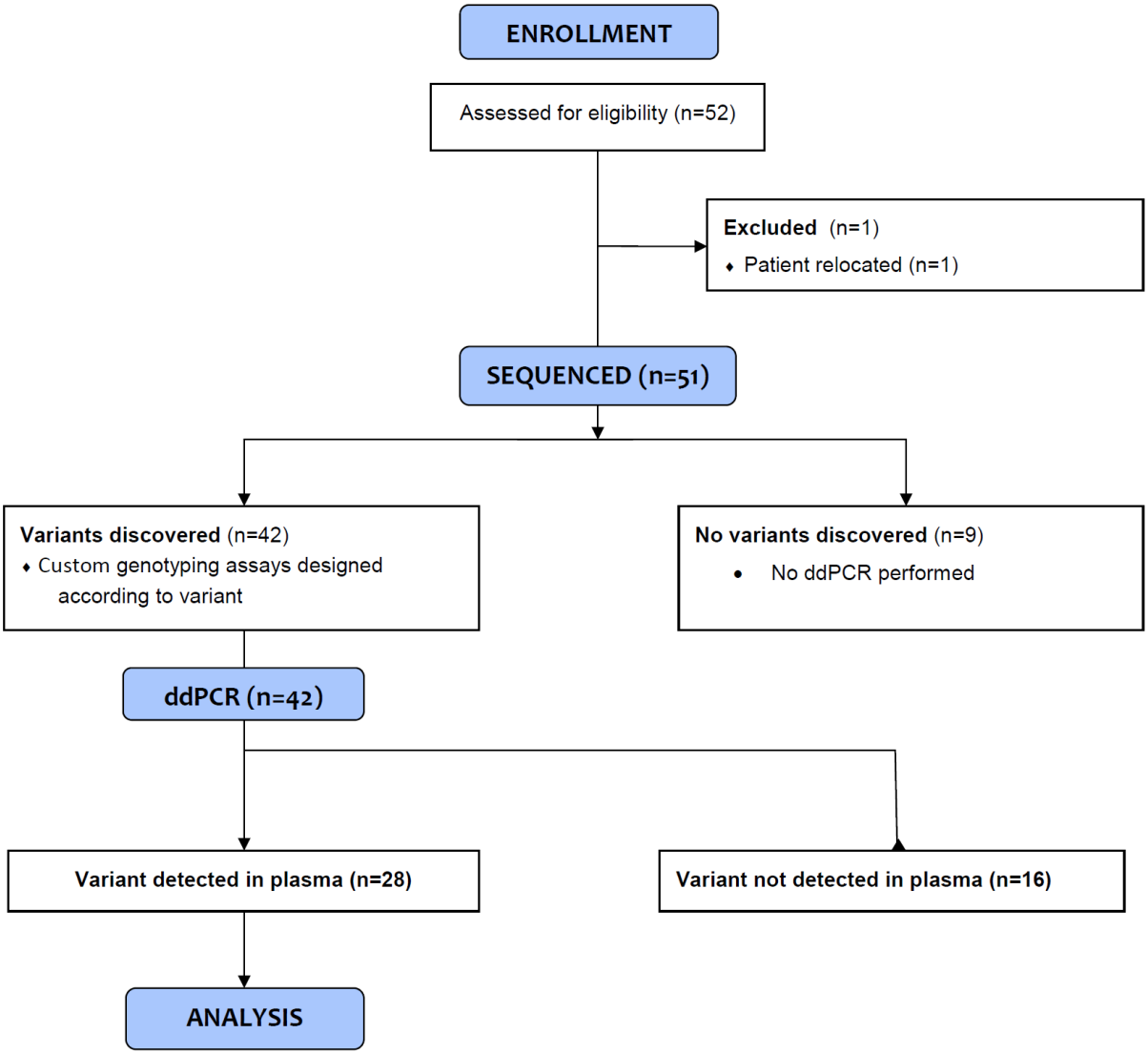

Figure 3-1. CONSORT diagram of study

The clinical characteristics of the patients are summarised in **Table 3-2** and shown in more detail in **Table 3-3.** Tissue blocks were retrieved for DNA extraction in 50 cases. In one case, blocks were not available and DNA extraction was performed from H&E stained slides. Primary tumour was used in 44 cases (85%) and metastatic tissue in 8 cases (15%). The time range from tumour tissue sampling to entry into the study was 1-30 years. A minimum of 3 plasma time-points per patient was aimed for, sampled every 3 months where possible, and a mean of 5 time-points was achieved. A total of 190 plasma samples were assayed.

Demographics	
Median Age, years (Range)	*61 (27 – 81)*
Female, n (%)	*26 (51)*
Tumour Histology	
Papillary, n (%)	*17 (34)*
Follicular	*15 (29)*
Poorly Differentiated	*3 (6)*
Anaplastic	*1 (2)*
Medullary	*15 (29)*
Disease Status at study entry	
No structural disease, n (%)	*6 (12)*
Local recurrence	*2 (4)*
Distant metastasis	*43 (84)*

Table 3-2. Summary of patient clinical characteristics

Trial ID	Sex	Age	TNM	Histology	Disease Status	Location of metastasis	Surgery	RIA	EBRT	TT
THY005	F	62	T3 N0 M1	P	Distant Met	Mediastinum, Lung	2008	X		
THY006	F	74	T4a Nx M1	P	Distant Met	Lung, Mediastinum, Bone	2001	X	X	
THY018	F	61	T3 N0 M0	P	No Macroscopic Disease		2016	X		
THY022	F	64	T2 Nx M0	P	Distant Met	Brain, Lung, Spine, Humerus	2002	X	X	X
THY024	M	69	T4 N1 M0	P	Distant Met	Thyroid Bed, Mediastinum, Lung	2004	X		
THY032	F	46	T3 N1b M0	P	Distant Met	Lung	2011	X		
THY035	M	74	T3 N1b M0	P	Distant Met	Lung, Bone, Adrenal, Mediastinal	2014	X	X	
THY038	F	47	T3 N0 M0	P	No Macroscopic Disease		2016	X		
THY039	F	71	T4b N0 M0	P	Distant Met	Bone, Chest wall	1998	X		
THY042	M	58	T2 N0 M0	P	Distant Met	Lung	2003	X		X
THY043	F	66	T1 N0 M0	P	Distant Met	Lung, Mediastinum	2010	X	X	X
THY046	F	57	T3 N1b M0	P	No Macroscopic Disease		2013	X		
THY049	M	56	T3 N1b M0	P	Local recurrence		2015	X	X	
THY050	M	27	T3 N1b M0	P	No Macroscopic Disease		2010	X		
THY051	M	60	T3 Nx M0	P	Distant Met	Mediastinum, Supraclavicular Fossa	2007	X	X	
THY007	M	48	T3 N0 M0	P	Distant Met	Brain	2010	X	X	
THY017	M	46	T4a N1b M1	P	Distant Met	Lung, Spine	2011	X	X	X
THY010	F	77	T3 Nx M0	F	Distant Met	Lung	2004	X		
THY011	F	30	T3 N0 M0	F	Distant Met	Lung, Brain, Bone	2013	X	X	X
THY012	M	72	T3 N0 M1	F	Distant Met	Bone	2008	X		
THY021	F	78	T3 N0 M0	F	Local recurrence		2015	X		
THY034	M	64	T2 N0 M0	F	Distant Met	Lung	2013	X		
THY036	M	56	T2 N1b M0	F	Distant Met	Lung	2011	X		
THY040	M	73	T1b Nx M1	F	Distant Met	Bone	2015	X	X	X
THY028	F	61	T1b N0 M1	F	Distant Met	Lung, Bone	2011	X	X	
THY008	F	67	T2 Nx M0	F	Distant Met	Liver, Lung	2010	X		X
THY013	F	81	T1b N0 M0	F	Distant Met	Thyroid Bed, Lung	2009	X	X	X
THY026	M	74	T4 N0 M0	F	Distant Met	Lung, Renal, Mediastinal	2007	X	X	X
THY027	M	51	T3 N1b M1	F	Distant Met	Lung	2013	X	X	
THY029	F	48	Tx Nx M1	F	Distant Met	Mediastinum, Spine, Lung	2008	X	X	X
THY044	F	49	T3 N1a M1	F	Distant Met	Mediastinum, Liver, Bone	2011	X	X	X
THY047	M	73	T3 N0 M0	F	Distant Met	Bone (rib)	2011	X	X	
THY031	F	70	T3 N0 M0	PD	Distant Met	Lung	2007	X	X	X
THY045	F	67	T3 N1a M1	PD	Distant Met	Lung, Mediastinum	2015	X	X	X
THY030	F	63	T2 N0 M0	PD	Distant Met	Liver, Brain	2007	X	X	X
THY033	M	66	T4b Nx M0	Ana	Distant Met	Bone, Mediastinum, Lung	nil		X	
THY001	M	54	T3 N1b M1	M	Distant Met	Lung, Liver, Mediastinum	2014			X
THY020	F	67	T2 N1b M0	M	Distant Met	Spine	2000		X	
THY023	F	33	T3 N1b M1	M	Distant Met	Liver, Mediastinum	2013			X
THY053	F	75	T2 N1b M0	M	No Macroscopic Disease		2005			
THY002	M	45	T4a N1b M1	M	Distant Met	Liver	2015		X	X
THY003	F	71	Tx N1a M0	M	Local recurrence		1987		X	
THY009	M	59	T4a N1b M0	M	Distant Met	Supraclavicular fossa	2009		X	X
THY014	M	68	Tx Nx M0	M	Distant Met	Mediastinal, pericardial	1983		X	X
THY015	M	41	Tx N1b M1	M	Distant Met	Lung, Liver, Brain	2005		X	X
THY016	M	54	T2 N1a M1	M	Distant Met	Liver, Spine, Lung	2002		X	X
THY019	M	28	T3 N1b M1	M	Distant Met	Mediastinal, Liver, Bone	2012		X	
THY037	M	67	T2 N1b N0	M	Distant Met	Bone, Liver, Lung, Mediastinal	1987		X	X
THY048	F	30	Tx N1b M1	M	Distant Met	Lung, Bone, Breast, Liver, Mediastinum	nil			X
THY041	M	32	T4a N1a M0	M (MEN 2a)	Distant Met	Bone, Orbit, Tracheal	2001		X	X
THY025	F	39	TxNxMx	M (MEN 2a)	Distant Met	Liver, Lung, Oesophageal, Spine	2003		X	X

Table 3-3. Clinical characteristics of enrolled patients. *TNM: Tumour Nodal Metastasis (7th Edition) staging at time of disease diagnosis. P: papillary thyroid cancer, F: follicular, PD: poorly differentiated, Ana: Anaplastic, M: medullary, MEN 2a: Multiple Endocrine Neoplasia syndrome type 2a. Disease status: disease status at time of enrolment into study. No macroscopic disease: no structural disease detectable on imaging with negligible conventional disease markers. Surgery: year of completion or total thyroidectomy. RIA: radioiodine ablation, cross indicates patient received. EBRT: external beam radiotherapy. TT: targeted therapy.*

3.4.2 Overview of variants detected in tumour

In total, 58 variants were discovered in 42 patient tumours. At least one variant was detected in 82% of the patient cohort. The gene capture panel and related analysis pipeline were not designed for the detection of structural variants (SVs): only a likely subset of existing variants are reported through this, and no SVs were detected. All reported variants, therefore, were either Single Nucleotide Variants (SNVs) or small insertion/deletions (Indels). A heatmap summarising discovered variants is shown in **Figure 3-2** and a full list is given in **Table 3-4**.

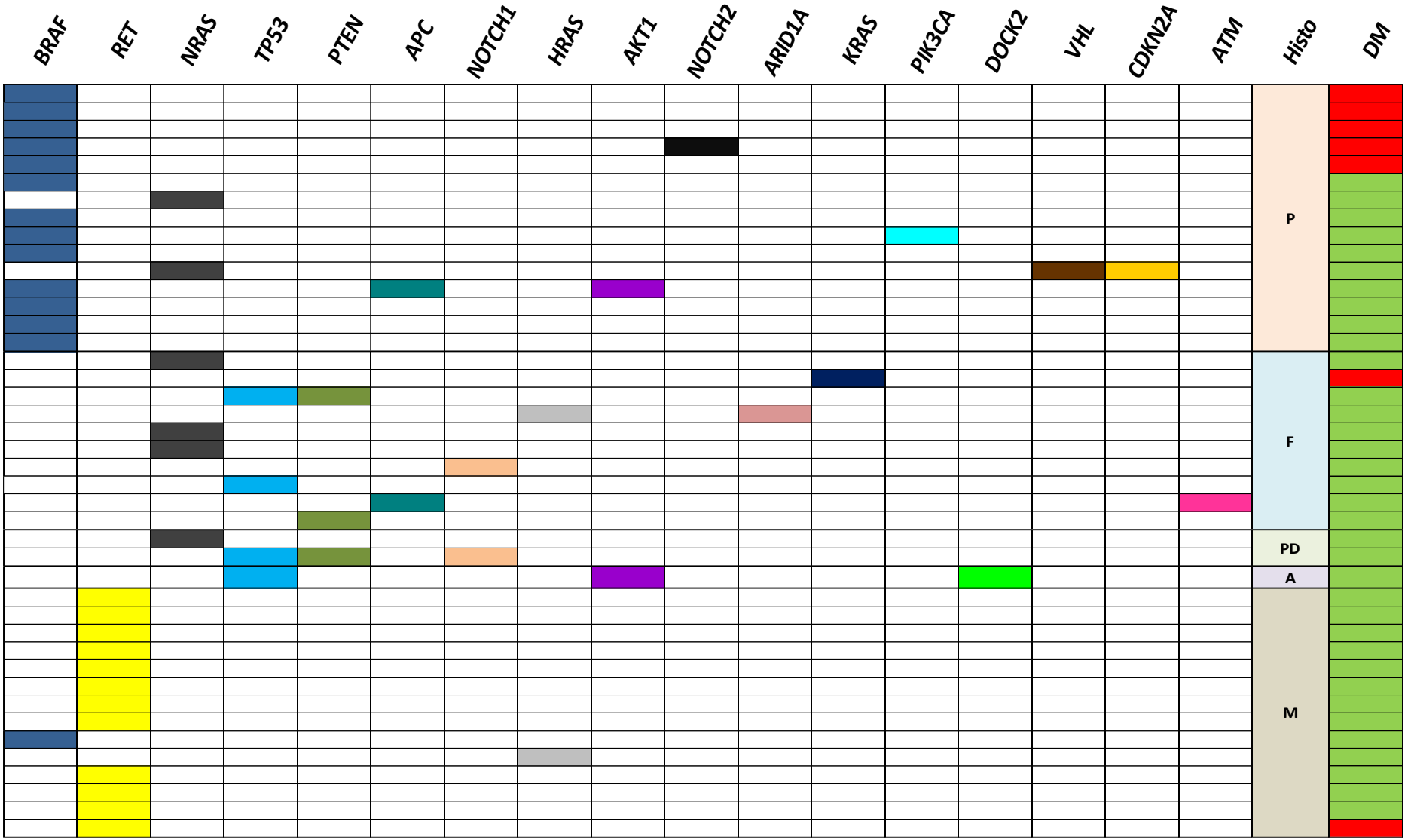

Figure 3-2. Mutation heat map from study tumour FFPE cohort*. Each row represents a different patient. Boxes are highlighted if a variant in the tumour was detected in the column's gene. Histo: Histology, P: Papillary, F: Follicular, PD: Poorly Differentiated, A: Anaplastic, M: Medullary, DM: Presence of distant metastasis at time of entry into study, green = yes, red = no*

Trial ID	GENE	PROTEIN	HGVS	VAF
THY033	*AKT1*	n/a	c.-21G>A	0.31
THY039	*AKT1*	p.Glu17Lys	c.49G>A	0.29
THY039	*APC*	p.Asp2133His	c.6397G>C	0.32
THY044	*APC*	p.Gln1916Ter	c.5746C>T	0.39
THY012	*ARID1A*	p.Gln601Ter	c.1800GC>TT	0.34
THY044	*ATM*	p.Phe213fs	c.634delT	0.71
THY005	*BRAF*	p.Val600Glu	c.1799T>A	Man
THY017	*BRAF*	p.Val600Glu	c.1799T>A	0.49
THY018	*BRAF*	p.Val600Glu	c.1799T>A	0.22
THY019	*BRAF*	p.Val600Glu	c.1799T>A	0.43
THY022	*BRAF*	p.Val600Glu	c.1799T>A	Man
THY024	*BRAF*	p.Val600Glu	c.1799T>A	0.07
THY032	*BRAF*	p.Val600Glu	c.1799T>A	0.31
THY038	*BRAF*	p.Val600Glu	c.1799T>A	0.36
THY039	*BRAF*	p.Val600Glu	c.1799T>A	0.38
THY042	*BRAF*	p.Val600Glu	c.1799T>A	0.4
THY043	*BRAF*	p.Val600Glu	c.1799T>A	Man
THY046	*BRAF*	p.Val600Glu	c.1799T>A	0.23
THY049	*BRAF*	p.Val600Glu	c.1799T>A	0.16
THY050	*BRAF*	p.Val600Glu	c.1799T>A	0.37
THY035	*CDKN2A*	p.Glu10Ter	c.28G>T	0.08
THY033	*DOCK2*	p.T459T	c.1377G>A	0.29
THY012	*HRAS*	p.Gln61Lys	c.181C>A	0.76
THY020	*HRAS*	p.Gln61Arg	c.182A>G	0.41
THY021	*KRAS*	p.Gln61Arg	c.182A>G	0.39
THY008	*NOTCH1*	p.Leu1932Phe	c.5796G>T	0.07
THY045	*NOTCH1*	p.Ala1135Val	c.3404C>T	0.37
THY049	*NOTCH2*	p.Ala21Thr	c.61G>A	0.12
THY006	*NRAS*	p.Gln61Arg	c.182A>G	0.45
THY010	*NRAS*	p.Gln61Lys	c.181C>A	0.49
THY028	*NRAS*	p.Gln61Arg	c.182A>G	0.54
THY030	*NRAS*	p.Gln61Arg	c.182A>G	0.48
THY035	*NRAS*	p.Gln61Arg	c.182A>G	0.43
THY036	*NRAS*	p.Gln61Arg	c.182A>G	0.39
THY024	*PIK3CA*	p.Tyr1021His	c.3061T>C	0.09
THY011	*PTEN*	p.Gln110Ter	c.328C>T	0.75
THY045	*PTEN*	p.Asn184fs	c.548dupA	0.64
THY047	*PTEN*	n/a	c.634+5G>C	0.62
THY001	*RET*	p.Met918Thr	c.2753T>C	0.45
THY002	*RET*	p.Glu632_Leu633del	c.1894_1899delGAGCTG	0.33
THY003	*RET*	p.Ala294Gly	c.881C>G	0.18
THY003	*RET*	p.Met918Thr	c.2753T>C	0.18
THY009	*RET*	p.Met918Thr	c.2753T>C	0.36

THY014	*RET*	p.Met918Thr	c.2753T>C	0.44
THY015	*RET*	p.Ala883Ser	c.2647GG>TT	0.24
THY016	*RET*	p.Met918Thr	c.2753T>C	0.48
THY023	*RET*	p.Met918Thr	c.2753T>C	0.39
THY037	*RET*	p.Ala883Ser	c.2647GC>TT	0.28
THY041	*RET*	p.Ser891Ala	c.2671T>G	0.21
THY048	*RET*	p.Met918Thr	c.2753T>C	0.67
THY053	*RET*		complex indel	0.4
THY011	*TP53*	p.Val272Met	c.814G>A	0.65
THY027	*TP53*	p.Arg337Cys	c.1009C>T	0.29
THY027	*TP53*	p.Asp281His	c.841G>C	0.06
THY027	*TP53*	p.Arg248Trp	c.742C>T	0.15
THY033	*TP53*	p.Arg158His	c.473G>A	0.28
THY045	*TP53*	p.Arg213Ter	c.637C>T	0.63
THY035	*VHL*	n/a	c.*50G>A	0.12

Table 3-4. List of detected variants from tumour sequencing. *VAF: variant allele frequency, HGVS: Human Genome Variation Society annotation. Man: low read number preventing pipeline from calling variant with VAF, but variant seen on manual curation and validated with dPCR*

3.4.3 Tumour variants detected by histological subtype

When analysed according to histological sub-type, the frequency of variants found in genes most often affected were broadly in line with rates published by The Cancer Genome Atlas (TCGA) (20) and the Catalogue of Somatic Mutations in Cancer (COSMIC) (128).

In the papillary cohort, 13 of the 21 (61%) detected variants were in *BRAF*, all of which were the c.1799T>A (p.V600E) variant. This compares similarly to COSMIC (55%) and TCGA (59%) numbers (**Figure 3-3**). Likewise, NRAS variants were detected at comparable rates. Mutations in BRAF were found to be mildly positively associated with PTC (p=0.07).

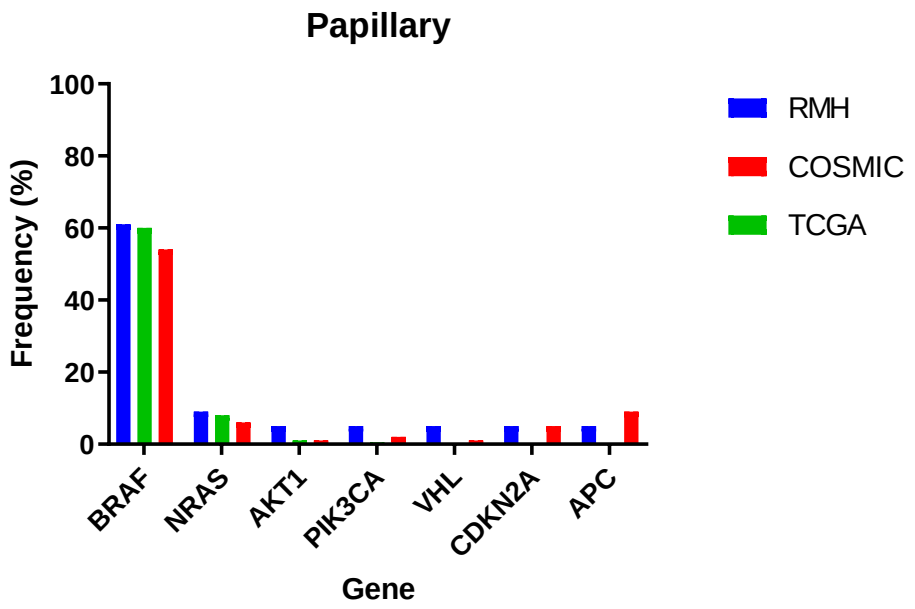

Figure 3-3. Frequency of gene variants detected in PTC tumour tissue, in comparison to published data. *RMH: Royal Marsden Hospital, rates obtained using GI2 panel. COSMIC & TCGA: published frequency on COSMIC database and TCGA respectively*

Figure 3-4 and **3-5** show the rates detected in follicular and medullary cancers respectively. Again, these were broadly similar to published data. The most frequently mutated gene in the FTC cohort was NRAS, contributing 20% of the total variants. Although not shown in **Figure 3-5**, the most commonly detected variant in the medullary cohort was *RET* c.2753T>C which accounted for 7 of 15 (46%). This is also in concordance with other published data series (28). A *RET* mutation was entirely specific for a medullary histology in this cohort, and was significantly positively associated with MTC ($p<0.001$)

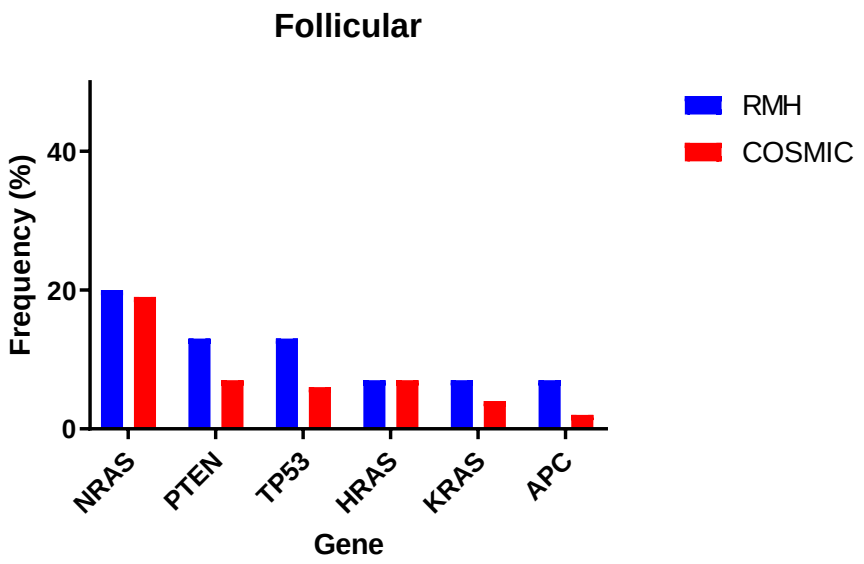

Figure 3-4. Frequency of gene variants detected in FTC tumours, in comparison to published data

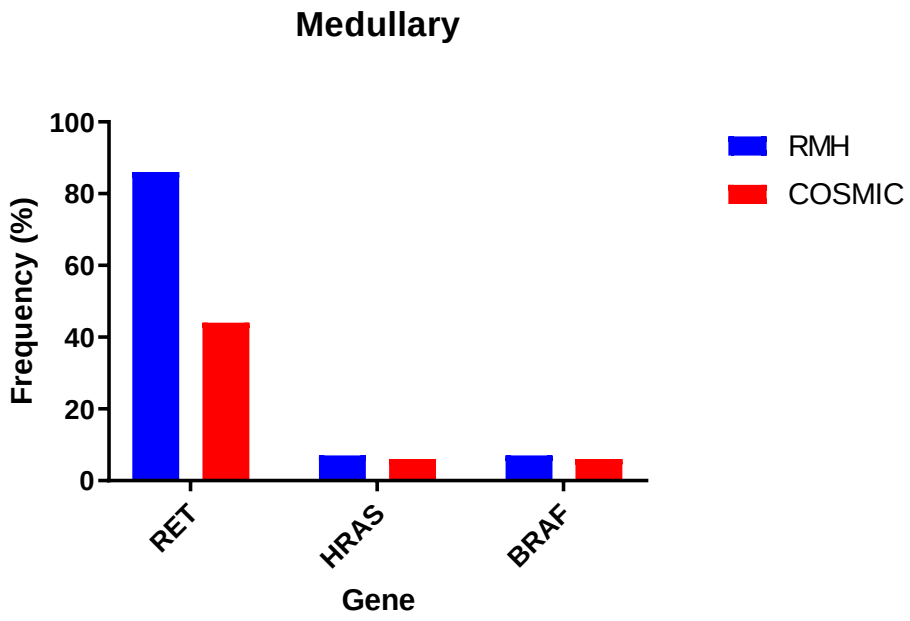

Figure 3-5. Frequency of gene variants detected in MTC tumours, in comparison to published data

3.4.4 Samples with ≥2 driver mutations

Nine tumour samples had two or more driver mutations, and all these patients had metastatic disease. Histological subtypes associated with more aggressive disease, such as Hürthle cell and PDTC were over-represented in this group.

3.4.5 Custom dPCR assay quality control

All custom dPCR assays used for ctDNA genotyping detected their target variant in tumour controls. There was very high correlation, R^2 = 0.92, between NGS and dPCR tumour allele frequency (**Figure 3-6**)

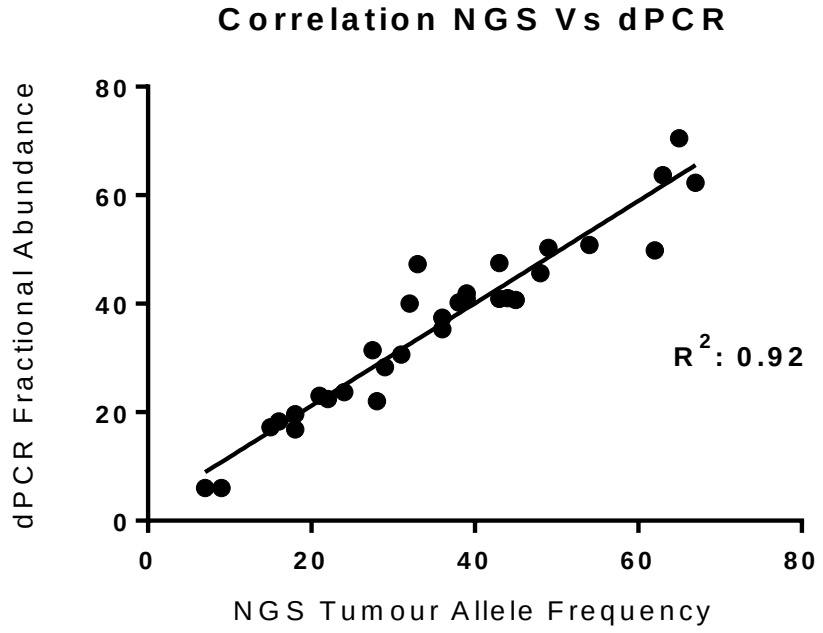

Figure 3-6. Scatter plot illustrating correlation between NGS and dPCR detected variant allele frequency.

3.4.6 Plasma ctDNA detection

The primary endpoint for the study was detection of plasma ctDNA. In this cohort ctDNA was detected in at least one plasma time point in 67% (n=28/42) of patients. Based on histology, the detection rate was higher in MTC compared to PTC and FTC (**Table 3-5**). However, five patients had no identifiable macroscopic disease throughout the study and negligible levels of conventional markers. If these patients were excluded from the analysis, the detection rates were more balanced across the histological subtypes and the overall detection rate increased to 76% (**Table 3-5**).The detection rate was highest (79%) in patients with metastatic disease compared to those with local recurrence (33%) or no macroscopic disease (0%). Detection of ctDNA was significantly higher in metastatic compared to non-metastatic disease ($p<0.001$). Median concentration of ctDNA was also significantly higher in metastatic disease compared to non-metastatic disease ($p<0.0001$).

Histology *(n)*	**Detection Rate**	***Adjusted***
Papillary *(15)*	53%	*73%*
Follicular *(10)*	60%	*60%*
Medullary *(14)*	79%	*85%*
PDTC *(2)*	100%	*100%*
ATC *(1)*	100%	*100%*
Total (42)	**67%**	***76%***

Table 3-5. Detection of ctDNA in patients at least one plasma time point. *Adjusted: detection rates of ctDNA adjusted for the 5 patients included in the cohort that had no macroscopic disease and negligible conventional markers throughout the study and in follow up afterwards.*

3.4.7 Correlation between ctDNA and conventional marker concentration

The secondary endpoint of the study was correlation of ctDNA with currently used biomarkers. Observationally, a clear overall trend was observed where trends in plasma ctDNA concentrations matched with Tg, calcitonin and CEA concentrations. This was borne out in the statistical analysis, which revealed a significant positive correlation between concentrations of ctDNA and those of conventional biomarkers (**Table 3-6**).

Conventional marker	***r***
ctDNA vs Thyroglobulin	0.53
ctDNA vs Calcitonin	0.62
ctDNA vs CEA	0.52

Table 3-6. Spearman rank correlation test of ctDNA concentration with conventional marker concentration. *P <0.0001 for all values.*

3.4.8 Correlation between ctDNA and disease progression on imaging

Median ctDNA concentrations were significantly higher (p=0.005) in patients with progressive disease (PD) on imaging compared to stable or responding disease (SD or PR). When change in ctDNA concentrations were modelled against disease progression on imaging, a small but positive effect was seen (p=0.1007).

3.4.9 Scenarios where ctDNA may offer superiority over conventional biomarkers

Serial serum analysis demonstrated that ctDNA markers may provide clinically important advantages over current plasma biomarkers in a number of cases. These include: detection of ctDNA in the absence of other markers, earlier changes in ctDNA levels reflective of disease status and more rapid detection of change in disease status when the patient is started on targeted therapies. These are discussed in turn below.

3.4.9.1 Detection of ctDNA in absence of other markers

Detection of ctDNA was possible in two patients that had no detectable levels of conventional biomarkers **(Figure 3-7**).

THY039 (**Figure 3-7a**) was a patient with PTC that had undetectable Tg levels due to anti-Tg antibodies. Two ctDNA variants were detectable in the plasma: *BRAF* c.1799T>A and *APC* c.6397G>C. A drop in ctDNA levels of both variants can be seen at month 6, which likely reflects an initial response of the disease to commencing targeted therapy (Sorafenib). This is followed by an increase in only the *BRAF* variant prior to disease progression from month 9 onwards. Disease progression was later confirmed on axial imaging. The *APC* variant remains low and is presumed to reflect a non-dominant subclone. Anti-Tg-Ig is seen to be trending downwards despite proven progressive disease.

THY 033 (**Figure 3-7b**) was a patient with anaplastic carcinoma and negligible or undetectable Tg levels, presumably due to severe tissue de-dedifferentiation. ctDNA levels are seen to increase prior to radiological disease progression, which is then confirmed at month 5 with axial imaging. The patient subsequently died of the disease. Both patients THY033 and THY039 had detectable ctDNA demonstrating upward trends before progression was detected with axial imaging.

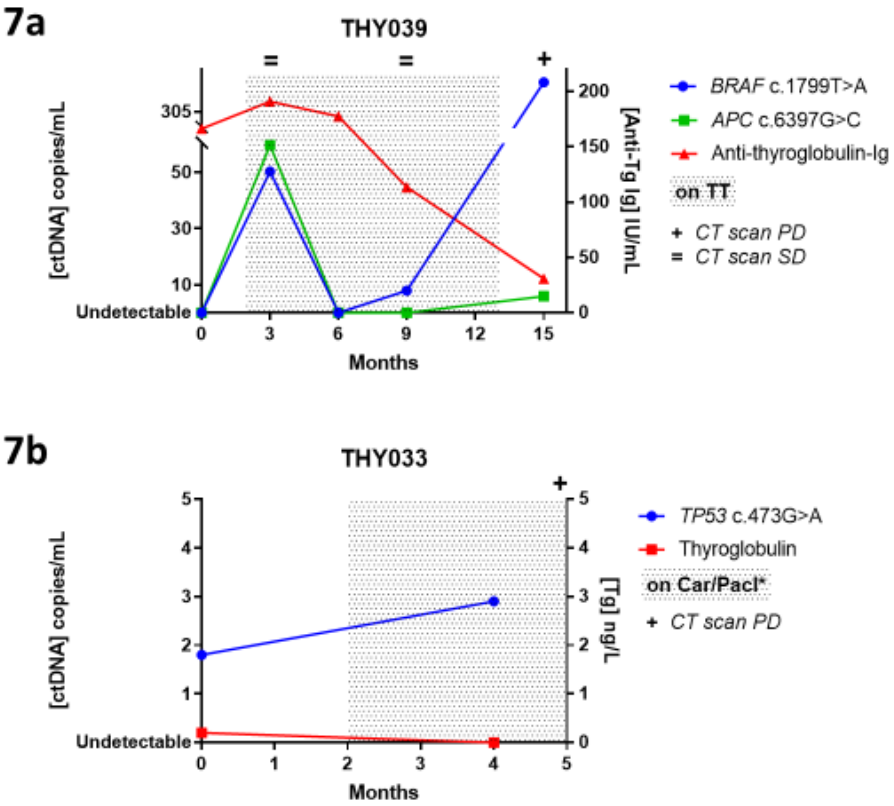

Figure 3-7. Detection of ctDNA when conventional markers absent*. TT: targeted therapy. *Car/Pacl: chemotherapy regimen of Paclitaxel and Carboplatin. CT Scan PD: Progressive Disease detected on CT axial imaging as per RECIST criteria. CT Scan SD: Stable disease. Shaded area represents period patient was on therapy.*

3.4.9.2 Earlier changes in ctDNA levels compared to conventional markers

In three patients with MTC who had evidence of progressive disease on imaging, an upward trend in ctDNA concentrations was noted earlier than in conventional markers (**Figures 3-8a** &3-8c). **Figure 3-8c** shows a patient with increasing levels of ctDNA, and subsequent calcitonin increase taken after the final ctDNA sampling. Axial imaging performed after the study period, at month 28 (not seen on the graph), showed progression of disease. In two patients with no conventional biomarkers available (**Figure 3-7**) an increase is seen prior to disease progression seen on imaging.

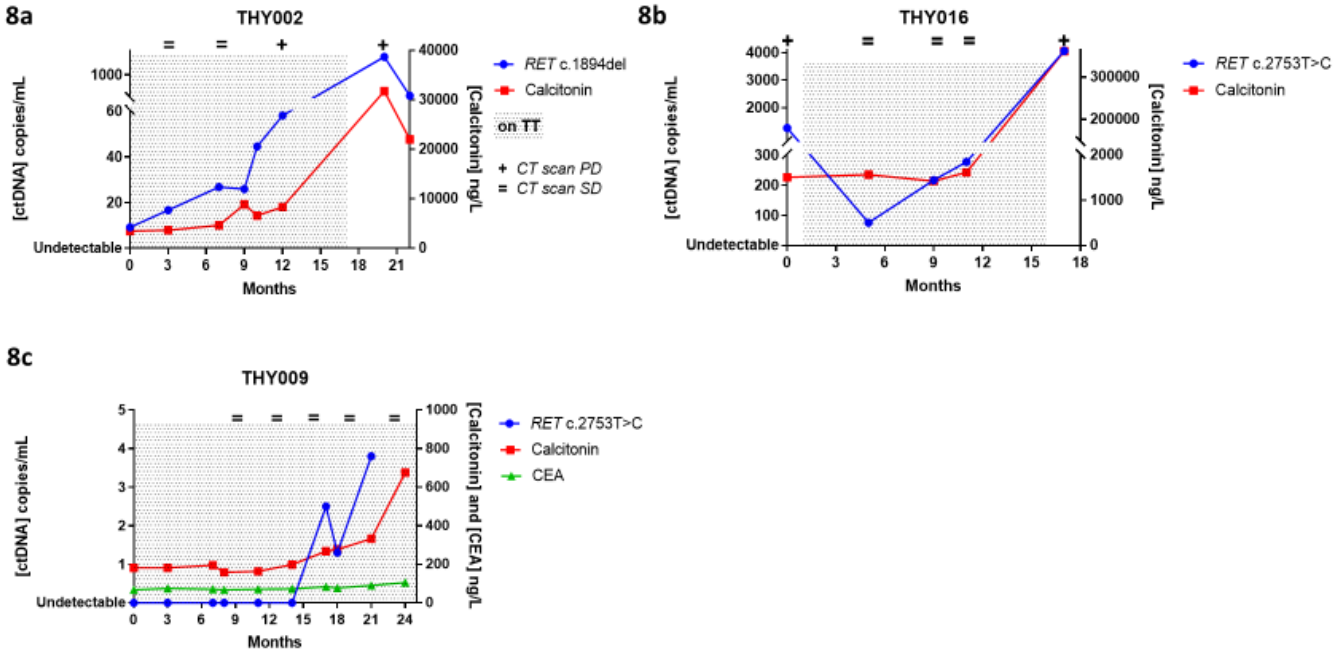

Figure 3-8. Earlier detection of progression in patients with MTC. *Three patients with MTC subtype showed an increase in ctDNA levels that pre-dated increases in conventional markers. Disease progression was later confirmed by CT imaging in all patients.*

3.4.9.3 Rapid detection of response to targeted therapies

In several patients on targeted therapy for progressive disease, changes were observed in ctDNA titres that were more marked or rapid than conventional markers (**Figures 3-9a:d**). Some patients showed more significant or earlier increases in ctDNA (**Figures 3-9a:c**), likely reflecting development of resistance to the targeted therapy. In contrast, THY028 (**Figure 3-9d**) demonstrated a large *decrease* in ctDNA levels after starting effective TKI therapy. Conventional biomarkers were unable to track these changes so responsively.

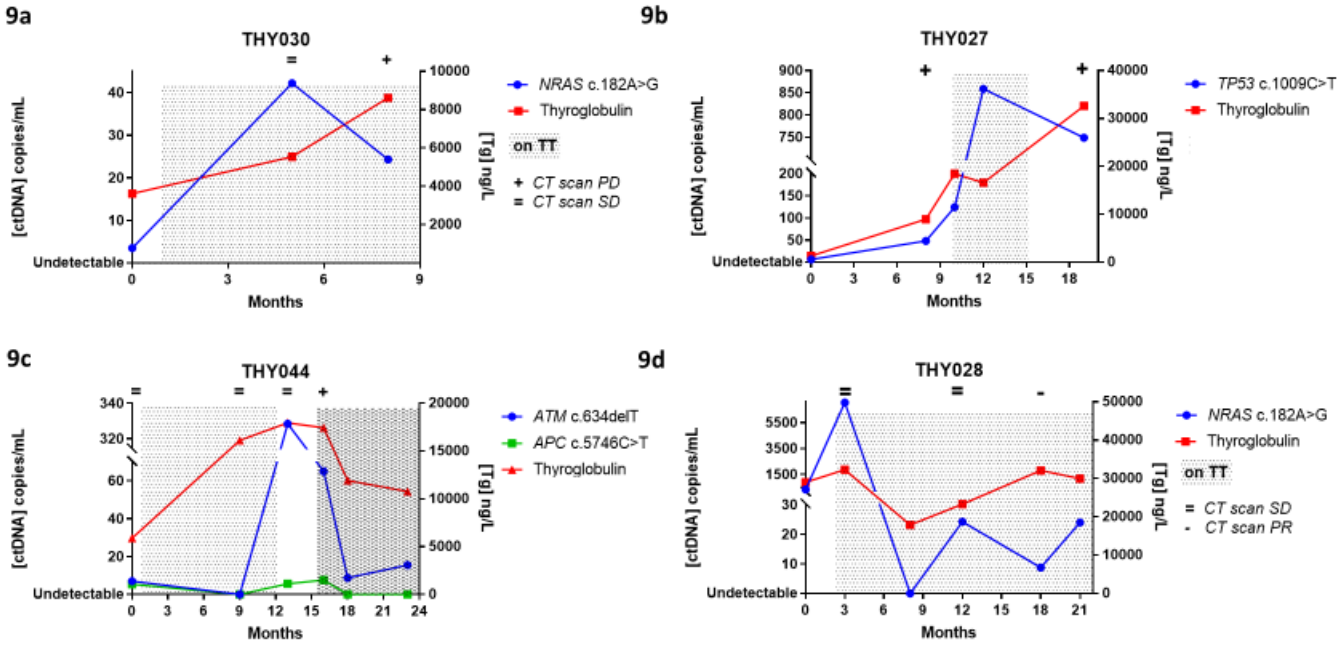

Figure 3-9. More responsive changes in ctDNA levels reflective of disease status. Figure 3-9a: *Ten-fold increase in ctDNA prior to detected progression on CT imaging. A similar increase is not seen in Tg levels, potentially offering more confident prediction of non-response to TKI.* ***Figure 3-9b:*** *Significant increase in plasma ctDNA levels at month 12, indicating non-response to TT, not mirrored in Tg until later.* ***Figure 3-9c:*** *Larger scale changes in ctDNA relative to Tg at month 12 and month 18 indicating disease progression followed by likely response to second TT.* ***Figure 3-9d:*** *marked ctDNA* drop *observed at month 8 following commencement of TT, indicating disease response later confirmed on imaging. The Tg levels were non-contributory in this case. CT scan PR: partial response noted on CT imaging as per RECIST criteria*

.

3.4.9.4 Subclonal evolution

In patients with two variants detected in the tumour which were tracked in the plasma, significant differences in the levels of one variant compared to the other were observed. In patient THY027 (**Figure 3-9b**) a *TP53* c.1009C>T variant was detectable in all samples, whereas the *TP53* c.742C>T variant, which is not shown on the graph, was not detected in any sample. In two other patients (**Figures 3-7a** and **3-9c**), both variants are initially seen to monitor disease burden. However, the subsequent progression of disease is tracked by only one variant. Differential levels of alternative variants may provide insight into tumour evolution.

3.5 Discussion

3.5.1 Tumour mutational analysis

The variants detected in the tumour tissue, and their frequency, were broadly in line with published data and those of the TCGA and COSMIC databases. In the COSMIC database, the gene most frequently observed to have variants in FTC is *TERT*. The targeted gene panel used in this study did not include this gene, so *TERT* mutations were not detectable in this study.

Nine patients had more than one driver mutation detected in the tumours. Five of these patients died of their disease during the study, giving a mortality rate of 55%. This compares to a 21% mortality rate for the whole cohort over the study duration. The presence of more than one driver mutation appears to be associated with poorer prognosis, as expected.

3.5.2 ctDNA analysis

The aim of this study was to investigate the potential role of a comprehensive, multi-variant mutational analysis using next generation sequencing and to report findings that provide support for the use of plasma ctDNA as a biomarker in thyroid cancer. Examining multiple genes in a variety of thyroid cancer histological subtypes, ctDNA was detected in 67% of patients in this study. Within the cohort of 42 patients in whom tumour variants were observed, 5 patients had no evidence of biochemical and structural disease throughout the study period. Therefore, the overall sensitivity for ctDNA detection was 76% (**Table 3-5)** and is comparable to rates obtained in other, more aggressive, cancers such as lung and breast (92).

Cote et al (126) recently evaluated the detection of a single variant, *RET* p.M918T, in the plasma of patients with MTC known to harbour this variant in their tumours. Their reported detection rate was only 32%. In the MTC cohort for this study, the detection rate was much higher (85%, adjusted). Both the use of one genetic variant as a biomarker and a single time point analysis in the Cote study are likely to have been responsible for their lower detection rates.

Interestingly, the rate of ctDNA detection was higher in patients with greater disease burden (metastatic versus non-metastatic disease), which is similar to results in other cancers. Also, the ctDNA concentrations in patients with metastatic disease were found to be significantly higher than those with non-metastatic disease, supporting the hypothesis that ctDNA levels are reflective of tumour burden. This may also explain the higher detection rates for our study in MTC patients (79%, unadjusted) all of whom had metastatic disease, compared to PTC

patients (53%, unadjusted). Further, in the more aggressive histological sub-types (ATC and PDTC), ctDNA was always detected, although numbers were low.

3.5.2.1 Correlation of ctDNA levels with conventional markers and imaging

Analysis of serial serum samples demonstrated a clear trend where detected plasma ctDNA concentrations correlated with the conventional biomarkers (**Table 3-6)**. Although significant ($p<0.0001$), higher correlation *r* values are unlikely to be observed given the increased variability of ctDNA concentrations compared to conventional markers, likely due to significantly shorter half-life. However, an advantageous consequence of this may be a more dynamic response to changes in disease status. Median ctDNA concentrations were found to be significantly higher in patients with disease progression on CT scans compared to those with stable or responding disease which supporting the view that concentration increases with increasing tumour activity. Using the linear modelling a small but positive association was found between changes in concentration and subsequent disease progression on imaging. Although this was not quite significant, p=0.1.007, this would support the view that increases in ctDNA concentration predict disease progression, and that with larger study numbers this would be statistically significant. In this sub-analysis there were only 18 patients that had progressive disease (PD) on imaging.

3.5.2.2 Areas where ctDNA may offer superiority over conventional markers

Although this study was primarily designed to evaluate ctDNA as a potential biomarker in thyroid cancer, this data provides early indications that ctDNA represents an important complementary assay and could offer several advantages over conventional markers.

ctDNA may prove to be an important biomarker of disease burden and response to treatment in patients when thyroglobulin cannot be used for monitoring, such as in patients with significant anti-thyroglobulin antibodies, and in de-differentiated thyroid cancers. In this study, one patient had undetectable thyroglobulin and significant anti-thyroglobulin antibodies at all time-points, yet ctDNA was detected for two variants in this patient's plasma and exhibited an upward trend in the *BRAF* variant coinciding with disease progression later confirmed on subsequent axial imaging (**Figure 3-7a**). In another patient with anaplastic carcinoma, the conventional biomarker, Tg, was not detectable, presumably due to tumour de-differentiation (**Figure 3-7b**). However, ctDNA was detectable at all time points and exhibited an upward trend consistent with disease progression which was later confirmed on axial imaging, and the patient subsequently died from the disease. Therefore, ctDNA may represent an important biomarker in anaplastic and de-differentiated thyroid cancers, particularly if effective therapies become available in the future.

In multiple patients, an upward trend in ctDNA was observed between 4 to 8 months earlier than the corresponding conventional biomarker (**Figure 3-8**), or a significantly more marked increase in levels (**Figures 3-9a-c**) was seen, allowing not only for earlier but also more

confident prediction of disease progression. These trends were later confirmed by axial imaging demonstrating disease progression.

Earlier detection was noted mainly in the MTC subtypes. This could be due to small study numbers observed during a relatively short time-frame, as not all patients progressed during the study period. Imaging scanning interval could also explain this, as blood samples were taken at regular three-monthly intervals, whereas imaging frequency was more variable, making it more difficult to ascertain objective disease progression for some patients. Disease biology may also play a role, as MTC is usually a more aggressive disease than DTC which may imply greater ctDNA release from increased cell turnover.

Assessing early biological response to targeted therapy and ascertaining when resistance develops in patients with thyroid cancer has traditionally been difficult. In this study, ctDNA levels appear to change more rapidly in response to a change in disease status than conventional markers and so may offer a more contemporaneous measure of response to treatment. In several patients, marked and rapid reduction of ctDNA levels was detected (**Figures 3-7a, 3-9c, 3-9d**) shortly after starting systemic targeted therapy. The more rapid responses of ctDNA levels in response to targeted therapy seen in this study would be useful in gauging early response to treatment. On the other hand, this might be of greater clinical value in detecting futility of therapy, thereby allowing prompt cessation of a potentially toxic but ineffective treatment. Further, development of resistance to treatment appears to be detected by ctDNA rises after the initial drop. It is evident that such changes in response cannot always to be monitored by conventional biomarkers (**Figures 3-7a, 3-8b, 3-9b-d**).

Prompt detection of resistance would allow the timely addition of alternative therapeutic agents to overcome newly-developing resistance. These more rapid changes in ctDNA concentrations may reflect significant differences in biomarker half-lives, which may be as low as 2 hours in ctDNA (48) and days to weeks for thyroglobulin (129, 130).

3.5.2.3 Subclonal evolution

Much has been published on the dynamic mutational landscape of cancers as a result of subclonal evolution (66-68), but data are sparse in thyroid cancer. In 3 patients, there was a substantial difference between the trends in the two tracked variants in the ctDNA, which was not reflective of the differences in tumour allele frequencies seen in the NGS results. This discrepancy may reflect subclonal evolution, as the tumour tissue samples often pre-dated the plasma samples by many years (**Table 3-3**). Discordant results obtained in the plasma may reflect one subclone emerging as more dominant than another. This has implications for patient therapy: knowledge of a dominant subclone may influence choice of intervention.

Chapter 4. Development and validation of a thyroid specific gene panel

4.1 Introduction

The study described in the previous chapter made use of a targeted gene panel, RMH GI2, that was designed primarily for colorectal cancers. This panel was chosen as similar studies had not previously been conducted for thyroid cancer in our unit, and a more thyroid-specific panel was therefore not available. Furthermore, there was a significant overlap between the mutations associated with cancers of both types. During the study the GI2 panel performed above expectations and detected variants in 82% of tumour samples.

However, whilst the panel was useful for the pilot study, it lacks published loci of interest that are relevant to thyroid cancer, e.g. the *TERT* promoter region. It also covers multiple unnecessary loci that are infrequently involved in thyroid cancer, increasing the amount of redundant data produced requiring analysis. Moreover, the GI2 panel uses capture-based target enrichment (TE) and was run on the Illumina NGS platform, necessitating larger amounts of DNA input and longer, more expensive library preparation compared to alternative techniques.

Given the positive findings of the ctDNA study, our department will move forward with further multi-centred studies involving molecular characterisation of thyroid cancers. For these studies it was felt that a thyroid-specific gene panel would be preferential.

Several proprietary thyroid-specific molecular tests that could be utilised already exist. These commercially available tests are primarily designed to provide increased diagnostic certainty for fine-needle aspiration cytology (FNAC) analysis of thyroid nodules, where 20-30% of samplings are inconclusive. Through molecular characterisation of genetic material obtained during FNAC sampling, these tests aim to provide further information to allow more definitive diagnosis of malignant or benign nodules when cytological examination has proven inconclusive. Examples of these tests include gene mutation tests based on NGS, such as the ThyroSeq V3 (CBLPath, Rye Brook, NY, USA), gene expression based tests such as the Afirma Gene Expression Classifier (Veracyte Inc., South San Francisco, CA, USA) and microRNA based tests such as the RosettaGX Reveal Thyroid Classifier (Rosetta Genomics Philadelphia, PA, USA). However, their use experimentally is limited due to several factors. These products are run as commercial enterprises, and thus cost is an inevitable factor: average medicare charge for ThyroSeq v.2 in 2018 was $4,056 (131). Furthermore, complete information on which genomic aberrations are covered by these assays is not easily accessible and there is no choice available in loci coverage to the client.

As such, it was felt that having a thyroid-specific, validated gene panel that was developed in-house would provide more versatility with regards to mutation coverage, reduce cost and turn-around time, as well as increasing diagnostic yield compared to the GI2 panel. Library preparation in amplicon-based target enrichment (TE) generally has a faster workflow than the capture-based TE used by the GI2 panel, and is also more cost-effective, hence why this method was chosen.

An ability to sequence directly from plasma would be advantageous clinically for a variety of reasons, such as detection of tumour heterogeneity and emergence of dominant subclones without recourse to tissue biopsy. Therefore, an important experimental consideration for this study was exploring the performance of this panel and associated workflow at sequencing ctDNA from plasma. Amplicon TE has a favourable performance at lower DNA concentrations compared to capture, further supporting the choice of this type of TE.

4.2 Hypothesis

A validated, thyroid-specific targeted gene panel would be non-inferior to previously used GI2 panel and could be used for future studies.

4.3 Aims

1. To design a targeted gene panel specific for thyroid cancer hotspots based on review of current literature
2. To validate the novel panel on thyroid tumour tissue previously sequenced with the GI2 panel
3. To explore usage of this novel panel for sequencing ctDNA in plasma to establish performance

4.4 Results

4.4.1 Novel panel design

A list of the exon coverage of the final panel design is shown in **Table 2-2**. A summary comparison with the GI2 panel is illustrated in **Figure 4-1**. The panel was designated "ThyMa" after ***Thy***roid Royal ***Ma***rsden Hospital.

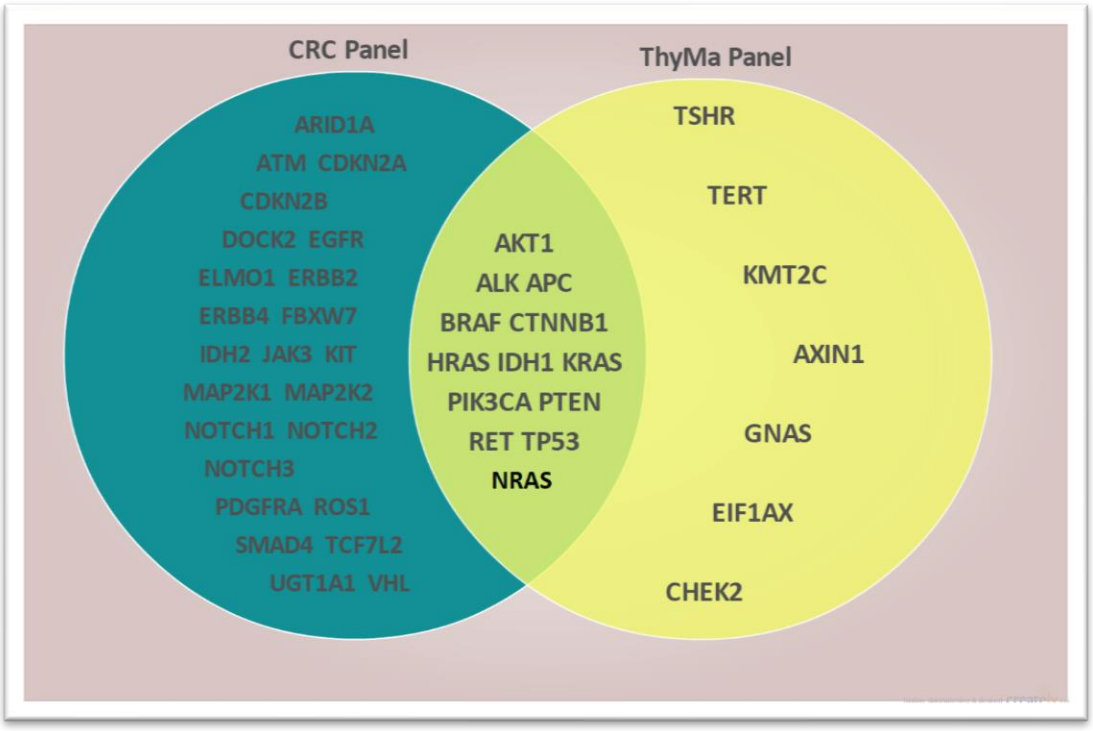

Figure 4-1. Comparison of genes covered by the 2 panels. *Left: RMH GI2 panel, Right: novel thyroid panel.*

4.4.2 Cost and coverage of novel panel

Cost of library preparation and sequencing was approximately £120 per sample. Total coverage of the ThyMa panel was 20.47 kb

4.4.3 Validation experiment

4.4.3.1 Samples Used

I aimed to sequence all the tumours (n=51) from the experiment in chapter 3 that were run with the GI2 panel. However, five samples had no DNA remaining. A further four samples failed quality control (QC), despite two attempts at library preparation (LP). This LP failure may have occurred due to low quality DNA: these samples were taken from FFPE extractions that were over 15 years old, had low concentrations and few reads were produced when they were run previously on the GI2/Illumina platform. In total, therefore, 42 tumour samples passed library preparation QC and were submitted for sequencing.

From the remaining samples in the previous experiment, a total of 26 plasma samples from 11 patients fulfilled the selection criteria. After LP QC, 22 samples remained and were sequenced (**Table 4-1**)

Sample ID	No. of samples	Passed LP QC
THY028	*2*	*2*
THY041	*3*	*1*
THY016	*5*	*5*
THY045	*1*	*1*
THY011	*1*	*1*
THY002	*3*	*2*
THY048	*2*	*2*
THY015	*5*	*4*
THY044	*2*	*2*
THY005	*1*	*1*
THY006	*1*	*1*
	TOTAL	***22***

Table 4-1. Numbers of plasma samples selected for sequencing by patient, and numbers passing library preparation quality control (LP QC). *No. of samples: number of plasma samples selected from that patient with sufficient sample remaining.*

The above FFPE and plasma samples, along with the two cell lines HTH-74 and TPC-1, were sequenced on a proton chip. The sequencing run generated 82,778,043 reads. The mean read length and mean depth across all samples were 96 bp and 3334x respectively. The median number of on target reads was 830,132 (80.6%) per sample.

4.4.3.2 FFPE Tumour sequencing results

4.4.3.2.1 ThyMa panel concordance with GI2 panel

The primary aim of the validation experiment was to ensure that the ThyMa panel was able to identify variants detected by the GI2 panel when covering the same loci. The ThyMa panel identified all variants detected with the GI2 panel when there was locus overlap, and so the panel performed satisfactorily. A summary of the results from the FFPE tumours is shown in **Table 4-2**. Six of the tumour samples had no prior calls using the GI2 panel, highlighted in italics in **Table 4-2,** and the ThyMa sequencing results from these samples are discussed in section ***4.4.3.2.2.*** Three tumour samples had calls with the GI2 panel that were in loci not covered by ThyMa, highlighted in bold, and the ThyMa results for these are discussed in ***4.4.3.2.1***.

Patient	Variant detected by GI2	Covered	Detected
THY001	*RET* c.2753T>C	Y	Y
THY002	*RET* c.1894_1899 del GAGCTG	Y	Y
THY003	*RET* c.2753T>C	Y	Y
THY006	*NRAS* c.182A>G	Y	Y
THY007	*no variants called*		
THY008	***NOTCH1* c.5796G>T**	**N**	
THY009	*RET* c.2753T>C	Y	Y
THY011	*TP53* c.814G>A	Y	Y
THY012	*HRAS* c.181C>A	Y	Y
THY013	*no variants called*		
THY014	*RET* c.2753T>C	Y	Y
THY015	*RET* c.2647GC>TT	Y	Y
THY016	*RET* c.2753T>C	Y	Y
THY017	*BRAF* c.1799T>A	Y	Y
THY018	*BRAF* c.1799T>A	Y	Y
THY019	*BRAF* c.1799T>A	Y	Y
THY020	*HRAS* c.182A>G	Y	Y
THY021	*KRAS* c.182A>G	Y	Y
THY022	*BRAF* c.1799T>A	Y	Y
THY023	*RET* c.2753T>C	Y	Y
THY024	*BRAF* c.1799T>A	Y	Y
THY025	*no variants called*		
THY026	*no variants called*		
THY027	*TP53* c.742C>T	Y	Y
THY028	*NRAS* c.182A>G	Y	Y
THY030	*NRAS* c.182A>G	Y	Y
THY031	*no variants called*		
THY032	*BRAF* c.1799T>A	Y	Y
THY035	*NRAS* c.182A>G	Y	Y
THY038	*BRAF* c.1799T>A	Y	Y
THY040	*no variants called*		
THY041	*RET* c.2671T>G	Y	Y
THY042	*BRAF* c.1799T>A	Y	Y
THY043	*BRAF* c.1799T>A	Y	Y
THY044	***ATM* c.634delT**	**N**	
THY045	TP53 c.637C>T	Y	Y
THY046	*BRAF* c.1799T>A	Y	Y
THY047	***PTEN* c.634+5G>C**	**N**	
THY048	*RET* c.2753T>C	Y	Y
THY049	*BRAF* c.1799T>A	Y	Y
THY050	*BRAF* c.1799T>A	Y	Y
THY053	*RET* c.complex indel	Y	Y

Table 4-2. Detection in tumour FFPE samples by ThyMa panel of variants previously called using GI2 panel. *Only 1 variant per sample shown when sample had more than 1 call. Covered: was GI2-called variant in a locus covered by the ThyMa panel. Detected: was the variant called using the ThyMa panel. Bold: patients with variant calls on GI2 panel in loci not covered by ThyMa panel. Italic: patients with no variant calls on GI2 panel.*

There was excellent correlation between tumour allele frequencies of detected variants across both panels, with a calculated Spearman rank *r* value of 0.91 (P <0.0001). Co-variance between the 2 allele frequencies is demonstrated in **Figure 4-2**.

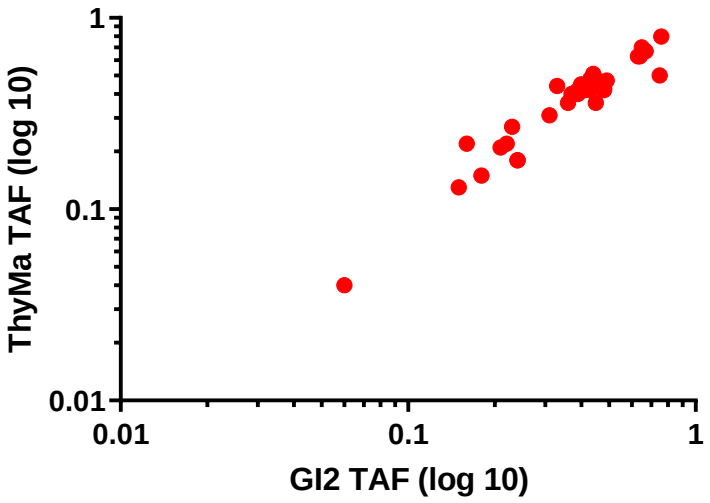

Figure 4-2. Co-variance of detected tumour allele frequencies (TAF) on Log 10 scale, between ThyMa and GI2 panel in tumour samples.

4.4.3.2.2 Novel variants detected by ThyMa panel

In total, 22 novel variants were called in the tumour samples that were not detected using the GI2 panel. Of these, 19 were in samples that had different variants called using the GI2 panel, and three were in samples that had no calls with the GI2 panel. These are described in turn below.

4.4.3.2.2.1 Novel ThyMa tumour variants detected in samples with *GI2 calls*

In tumour samples that had variants identified with the GI2 panel, 19 new variants were called with the ThyMa panel. These are summarised in **Table 4-3**. Three of these calls were in two patients, THY044 and THY047, whose GI2 detected variants were in loci that were not covered by the ThyMa panel. These three calls were important, as otherwise no putative driver mutations would have been called in these samples using the ThyMa panel. The variants were *TERT* c.1-124C>T and *KMT2C* c.1373delA in THY044, and *CHEK2* c.591delA in THY047. Only one patient that had no calls with ThyMa but did with GI2 panel: THY008, where a *NOTCH1* c.5796G>T variant was found. A significant number of the new calls (n=6, 31%) were in the *TERT* promoter region. The TERT variants, and the variants identified in THY044 and THY047, were validated with digital PCR assays, and these are discussed in section ***4.4.3.2.3***

Patient	Variant	TAF	Effect
THY006	TERT c.1-124C>T	*0.61*	Upstream promoter*
THY012	TERT c.1-124C>T	*0.56*	Upstream promoter*
THY014	AXIN1 c.2178delC	*0.27*	Frameshift
THY021	EIF1AX c.38G>T	*0.4*	Non-synonymous
THY022	TERT c.1-124C>T	*0.53*	Upstream promoter*
THY027	APC c.3743C>T	*0.16*	Non-synonymous
THY028	AXIN1 c.2209C>T	*0.21*	Non-synonymous
THY028	TERT c.1-124C>T	*0.83*	Upstream promoter§
THY035	KMT2C c.967A>G	*0.18*	Non-synonymous
THY035	AXIN1 c.17_18insC	*0.24*	Frameshift
THY041	APC c.1355delT	*0.07*	Frameshift
THY042	TERT c.1-146C>T	*0.39*	Upstream promoter
THY043	ALK c.2755G>A	*0.13*	Non-synonymous
THY044	**TERT c.1-124C>T**	***0.57***	**Upstream promoter***
THY044	**KMT2C c.1373delA**	***0.07***	**Frameshift**
THY045	TP53 c.C637T	*0.63*	Stop-gain
THY047	**CHEK2 c.591delA**	***0.88***	**Frameshift***
THY049	TERT c.1-124C>T	*0.33*	Upstream promoter*
THY053	RET c.1888T>G	*0.45*	Non-synonymous
THY006	TERT c.1-124C>T	*0.61*	Upstream promoter

Table 4-3. Summary of ThyMa detected variants not detected with GI2 panel, in samples that had other GI2 called variants. *Variant: description of variant according to Human Genome Variation Society (HGVS) nomenclature. TAF: tumour allele frequency. Bold: patients with GI2 calls in loci not covered by ThyMa panel. *: variant confirmed in tumour with dPCR validation. § insufficient DNA remaining for dPCR validation.*

4.4.3.2.2.2 Novel ThyMa tumour variants detected in samples without *GI2 calls*

In the previous study (chapter 3), nine tumour samples had no variant calls using the GI2 panel. Six of these were also sequenced with the ThyMa panel. The remaining three were not run due either to insufficient sample or failure of library preparation QC.

Four new calls were made in these six tumour samples. Three of these novel variants were *TERT* c.1-124C>T, and these were subsequently validated using dPCR (see section ***4.4.3.2.3.1***). One patient, THY025, was found to have a *RET* c.1858T>C variant. This is in a locus covered

by the GI2 panel and therefore should have been detected by the GI2 panel. The BAM files from the GI2 run were checked retrospectively and this variant was found in both the tumour and buffy coat data. This is a patient with MTC that is known to have Multiple Endocrine Neoplasia (MEN) Type 2A, a hereditary condition. This variant had been filtered by the pipeline used with the GI2 panel due its presence in non-tumour cells, indicating a germ-line mutation and is consistent with hereditary MTC. Two patients, THY007 and THY031, were successfully run with the ThyMa panel, but had no calls on either panel. A summary of the results from these patients, along with the histology and disease status on study entry, is shown in **Table 4-4.**

Patient	GI2 run	ThyMa run	ThyMa call	Hist	Status
THY007	✓	✓	***no***	PTC	DM
THY013	✓	✓	*TERT* c.1-124C>T	FTC	DM
THY025[¶]	✓	✓	*RET* c.1858T>C	MTC	DM
THY026	✓	✓	*TERT* c.1-124C>T	FTC	DM
THY029	✓	*no sample*		FTC	DM
THY031	✓	✓	***no***	PDTC	DM
THY034	✓	*no sample*		FTC	DM
THY040	✓	✓	*TERT* c.1-124C>T	FTC	DM
THY051	✓	*fail LP*		PTC	DM

Table 4-4. List of tumour samples with no calls with GI2 panel and ThyMa results*. GI2 and ThyMa run: sample run with respective panel. ThyMa call: variant called using ThyMa panel when detected. No: no variant detected. Shaded cells: no call as sample not run with ThyMa panel. ¶:Patient where variant detected by filtered by GI2 pipeline due to germline. Status: disease status. DM: metastatic disease.*

4.4.3.2.3 Validation of novel variants detected with ThyMa panel with dPCR

Novel variants discovered by the ThyMa panel, that were in loci *not* covered by the GI2 panel in samples that had either no GI2 call or a GI2 call in a non-covered loci by the ThyMa panel were confirmed with digital PCR assays. This included variants in the *TERT* promoter region,

KMT2C and *CHEK2* genes. The results of each are discussed in turn below. Other novel variants were not validated as these were in samples that already had concordant calls between GI2 and ThyMa panels and were thus deemed less relevant in the context of a limited budget.

4.4.3.2.3.1 TERTp variants

As reliable sequencing of the *TERT* promoter region was felt to be critical to the success of the panel, validation of these calls was attempted to ensure positive results were accurate. DNA samples extracted from two cell lines known to be positive for the *TERT* c.1-124C>T were also sequenced to help with validation. Sequencing of HTH-74 and TPC-1 both resulted in positive variants calls, offering confirmatory positive controls.

Overall, ten FFPE samples had calls in the *TERT* promoter region: nine were *TERT* c.1-124C>T and one was *TERT* c.1-146C>T. This included the three samples mentioned in section **4.4.3.2.2.2** that had no calls with the GI2 panel. Eight of these calls were validated using dPCR. The two samples that were not confirmed were due either to insufficient sample remaining or failure to design a functional *TERT* c.1-146C>T TaqMan assay for dPCR testing. The results are summarised in **Table 4-5.**

Patient	Variant	TAF	Hist	Status	dPCR
THY006	*TERT* c.1-124C>T	0.61	Papillary	DM	✓
THY012	*TERT* c.1-124C>T	0.56	Follicular	DM	✓
THY013	*TERT* c.1-124C>T	0.44	Follicular	DM	✓
THY022	*TERT* c.1-124C>T	0.54	Papillary	DM	✓
THY026	*TERT* c.1-124C>T	0.50	Follicular	DM	✓
THY040	*TERT* c.1-124C>T	0.28	Follicular	DM	✓
THY044	*TERT* c.1-124C>T	0.57	Follicular	DM	✓
THY049	*TERT* c.1-124C>T	0.34	Papillary	LR	✓
THY028	*TERT* c.1-124C>T	0.83	Follicular	DM	***no DNA***
THY042	*TERT* c.1-146C>T	0.39	Papillary	DM	***assay fail***

Table 4-5. Summary results of dPCR validation of TERTp variants in FFPE samples*. LR: Local recurrence. dPCR valid: variant confirmed detected on dPCR. No DNA: insufficient sample remaining. Assay fail: failure to design dPCR assay.*

There was good correlation between the tumour allele frequencies detected by sequencing with the ThyMa panel compared to dPCR (r = 0.84, P = 0.01). Co-variance is shown in **Figure 4-3**.

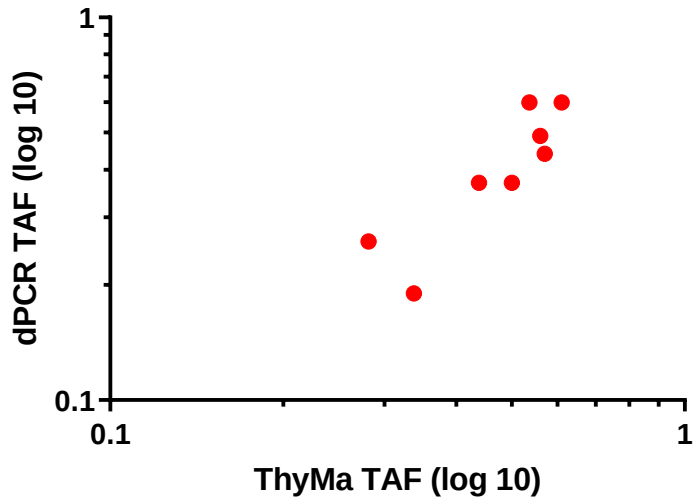

Figure 4-3. Co-variance of detected tumour allele frequencies for variant TERT c.1-124C>T in tumour samples between NGS ThyMa sequencing and dPCR assay.

Detection rates of *TERT*p variants were markedly higher in this cohort than rates published as part of the COSMIC database. Expected and actual results per histology are shown in **Table 4-6**. This is discussed further in section **4.5.**

Subtype	*Expected*	*Called*
FTC	16%	40% *(6 of 15)*
PTC	11%	23% *(4 of 17)*

Table 4-6. Percentage of FFPE samples expected to have a TERTp variant (based on data derived from COSMIC) compared to those called through sequencing with ThyMa panel.

4.4.3.2.3.2 KMT2C variant

The tumour sample from patient THY044 was found to have the frameshift mutation *KMT2C* c.1373delA. This gene codes for a methyltransferase enzyme involved in histone methylation. As this was a new variant discovered in a sample that had GI2 identified variants in loci not covered by the ThyMa panel, it was felt important to attempt validation. TaqMan genotyping probes were designed for this variant. Although good separation of droplets was seen with the wild-type DNA, suggestive of a functional assay, no positives were seen with the mutant (**Figure 4-4** and **Figure 4-5**). The NGS findings were therefore not confirmed with dPCR.

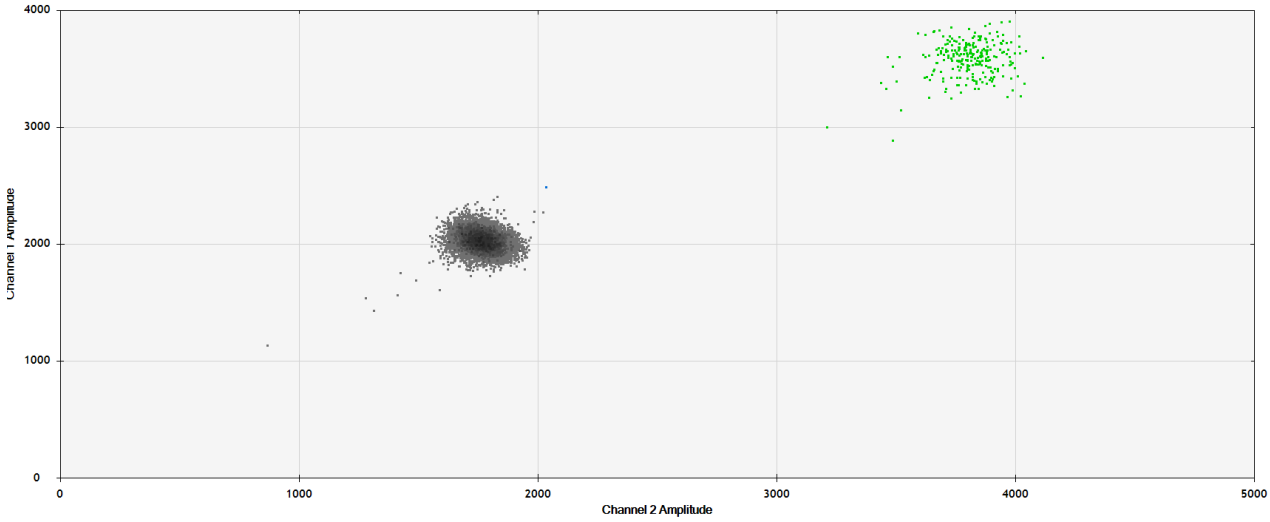

Figure 4-4. dPCR results of wild-type (WT) control (MCF-12A cell line) with KMT2C c.1373delA assay. Good *droplet separation between WT positives (green) and negative(grey) droplets are seen*

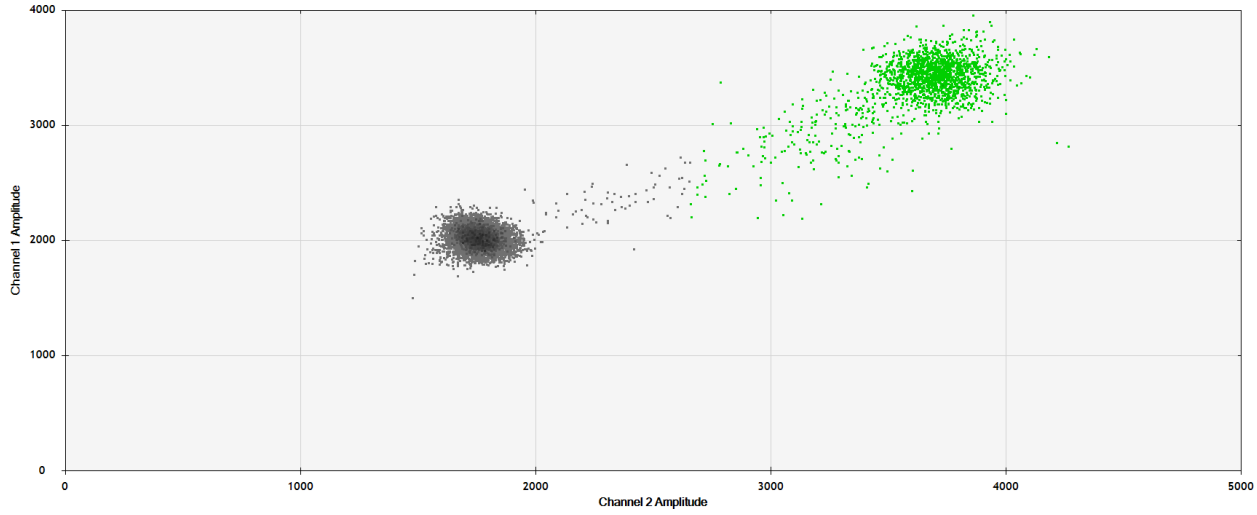

Figure 4-5. dPCR results of THY044 tumour sample, in which KMT2C c.1373 delA was detected on sequencing with ThyMa panel. *Only WT positive and negative droplets are seen, no mutants.*

4.4.3.2.3.3 CHEK2 variant

Checkpoint kinase 2 (*CHEK2*) is a tumour suppressor gene: it acts to halt cell cycle progression in response to DNA damage. The tumour sample from THY047 had a positive call for a *CHEK2*

c.591delA variant, a frameshift mutation. The detected tumour allele frequency was 0.88. Again, after assay design, the sample was submitted for confirmatory detection with digital PCR.

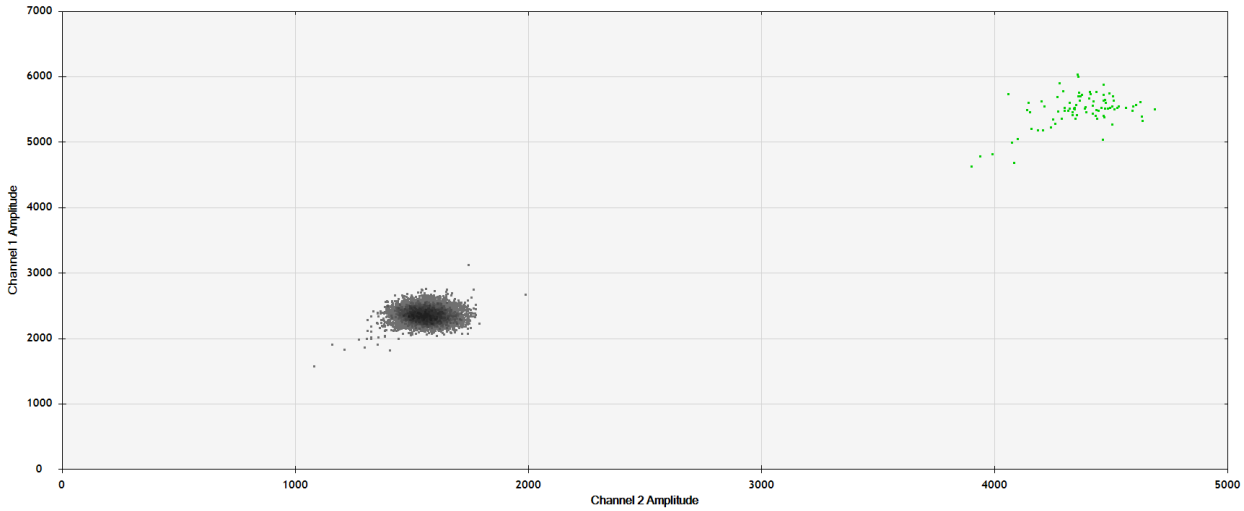

***Figure 4-6. dPCR results of wild-type control (MCF-12A cell line) with CHEK2 c.591delA. Good** droplet separation between WT positives and negative droplets are seen*

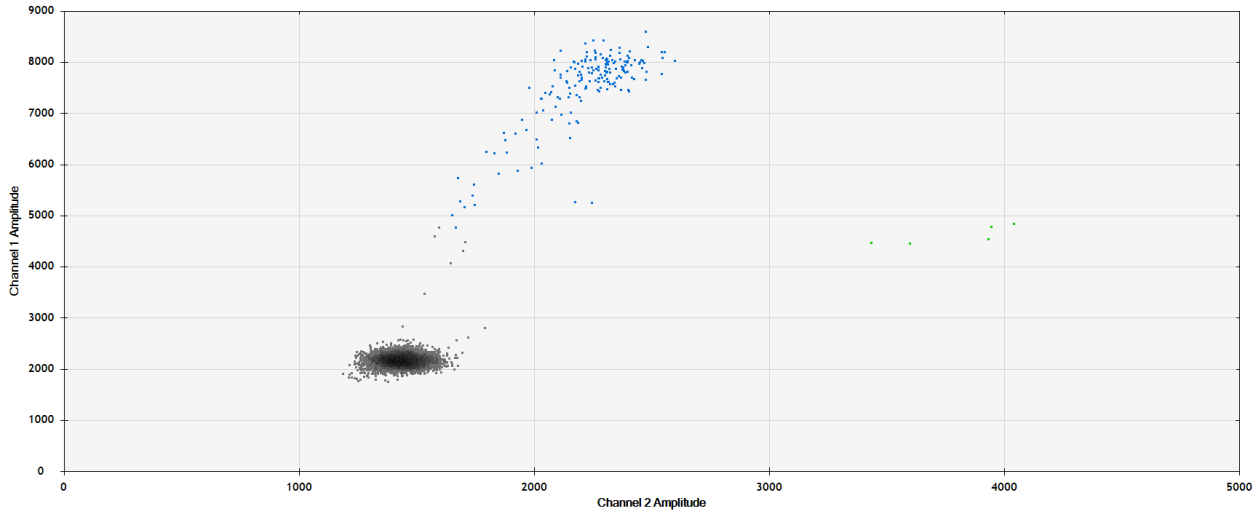

***Figure 4-7. dPCR results of THY044 tumour sample in which CHEK2 c.591delA was detected on sequencing with ThyMa panel.** Predominantly blue drops are seen, representing mutant CHEK2 detection, and a few scattered green droplets (CHEK2 WT)*

Figure 4-6 and **Figure 4-7** are shown above to illustrate a positive result using dPCR. **Figure 4-6** depicts the results from the wild-type control, where no mutant CHEK2 was detected. The cluster of green droplets represent detected wild-type DNA. This compares to **Figure 4-7**, where both green and blue droplets are seen, indicating mutant DNA detection. The dPCR detected fractional abundance/tumour allele frequency of the *CHEK2* c.591delA variant was 0.95, thus validating the ThyMa NGS findings.

4.4.3.2.3.4 Detection of CHEK2 c.591delA variant in plasma

In the previous study (chapter 3) a PTEN variant had been called in the tumour of patient THY047 using the GI2 panel. Five plasma time points from that patient were tested for ctDNA with dPCR, but the PTEN variant was not detected in any of the samples. Given the discovery of the *CHEK2* c.591delA variant in the tumour using the ThyMa panel, the same plasma samples were re-tested using the *CHEK2* variant dPCR assay. Of the five plasma samples that were assayed again, four tested positive for this variant.

4.4.3.2.3.5 ThyMa vs GI2 panel variant detection rate in tumour samples

The ThyMa panel had a higher number of patients with variant calls than the GI2 panel. At least one variant was detected in 39 of the 42 (93%) tumour samples run with the ThyMa panel. For the same tumours, the GI2 panel detected a variant in 36 (86%) samples only.

4.4.3.3 Plasma sequencing results

4.4.3.3.1 Detection of known tumour variants in plasma

The 22 plasma samples sequenced were obtained from 10 patients. The first analysis looked at whether the sequencing data from these plasma samples detected the variant previously found in them with dPCR. Two plasma samples from patient THY044 were excluded from this part of the analysis as the dPCR-tracked variant was in a loci not covered by the ThyMa panel. Thus 20 plasma samples were included, and the dPCR-tracked variant was detected in 13 (65%) samples. A summary of the results is shown in **Table 4-7**.

Patient	Variant	Sample	[cfDNA]	dPCR AF	Detect	ThyMa AF
THY002	*RET* c.1894delGAGCTG	*05/12/16*	0.23	2.0	*n*	
		06/02/17	0.30	1.8	*n*	
THY005	*BRAF* c.1799T>A	*16/08/17*	0.20	2.6	*n*	
THY006	*NRAS* c.182A>G	*27/04/16*	1.42	2.6	Y	2.6
THY011	*TP53* c.814G>A	*01/06/16*	1.47	19.8	Y	19.1
THY015	*RET* c.2647_2648GC>TT	*07/12/16*	0.29	3.8	Y	1.4
		08/02/17	0.59	2.1	*n*	
		21/06/17	0.67	2.4	*n*	
		13/09/17	0.50	7.1	Y	2.3
THY016	*RET* c.2753T>C	*06/07/16*	2.28	7.7	Y	5.3
		02/12/16	0.56	1.7	Y	1.4
		26/04/17	6.80	2.8	Y	1.1
		21/06/17	0.94	3.8	Y	2.2
		06/12/17	8.12	9.1	Y	4.5
THY028	*NRAS* c.182A>G	*09/03/16*	1.03	5.9	Y	3.5
		01/06/16	10.30	18.2	Y	14.2
THY041	*RET* c.2671T>G	*01/06/16*	2.26	9.1	Y	9.3
THY045	*TP53* c.637C>T	*23/01/17*	0.68	2.7	Y	2.6
THY048	*RET* c.2753T>C	*01/04/16*	0.32	7.2	*n*	
		16/11/16	0.43	10.1	*n*	

Table 4-7. ThyMa sequencing results from plasma samples: detection of known variants previously detected in plasma by digital PCR. *Sample: date of patient plasma sampling. [cfDNA]: concentration of cell free DNA concentration in plasma (ng/µL). dPCR AF: allele frequency (as %) of variant detected in plasma by dPCR. Detect: was variant seen in plasma sequencing data, Y: yes, n:no. ThyMa AF: allele frequency of variant when detected.*

When a variant was identified through sequencing, there was good correlation with variant allele frequencies detected with dPCR (r = 0.81, P = 0.001). This is visualised in **Figure 4-8**.

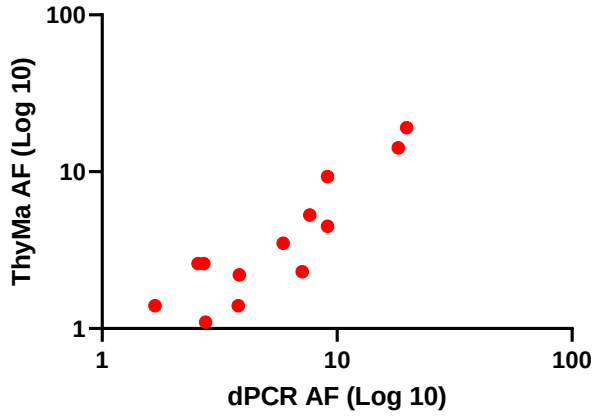

Figure 4-8. Co-variance of detected variant allele frequencies in plasma from ThyMa NGS sequencing compared to dPCR*. AF: variant allele frequency.*

When the plasma samples were ranked according to the dPCR detected variant allele fraction and DNA concentration, an observable trend was seen. Detection rates at higher TAF and DNA concentrations appear greater (**Table 4-8**), as expected. Crudely, variants were detected when allele fractions were greater than 2.60% and cfDNA concentrations were above 0.67 ng/µL.

Sample	Variant	AF	Det	Sample	Variant	[DNA]	Det
02/12/2016	*RET* c.2753T>C	1.7	Y	*16/08/2017*	*BRAF* c.1799T>A	0.20	n
06/02/2017	*RET* c.1894del	1.8	n	*05/12/2016*	*RET* c.1894del	0.23	n
05/12/2016	*RET* c.1894del	2.0	n	*07/12/2016*	*RET* c.2647_2648	0.29	Y
08/02/2017	*RET* c.2647_2648	2.1	n	*06/02/2017*	*RET* c.1894del	0.30	n
21/06/2017	*RET* c.2647_2648	2.4	n	*01/04/2016*	*RET* c.2753T>C	0.32	n
16/08/2017	*BRAF* c.1799T>A	2.6	Y	*16/11/2016*	*RET* c.2753T>C	0.43	n
27/04/2016	*NRAS* c.182A>G	2.6	n	*13/09/2017*	*RET* c.2647_2648	0.50	Y
23/01/2017	*TP53* c.637C>T	2.7	Y	*02/12/2016*	*RET* c.2753T>C	0.56	Y
26/04/2017	*RET* c.2753T>C	2.8	Y	*08/02/2017*	*RET* c.2647_2648	0.59	n
07/12/2016	*RET* c.2647_2648	3.8	Y	*21/06/2017*	RET c.2647_2648	0.67	n
21/06/2017	*RET* c.2753T>C	3.8	Y	*23/01/2017*	*TP53* c.637C>T	0.68	Y
09/03/2016	*NRAS* c.182A>G	5.9	Y	*21/06/2017*	*RET* c.2753T>C	0.94	Y
13/09/2017	*RET* c.2647_2648	7.1	Y	*09/03/2016*	*NRAS* c.182A>G	1.03	Y
01/04/2016	*RET* c.2753T>C	7.2	n*	*27/04/2016*	*NRAS* c.182A>G	1.42	Y
06/07/2016	*RET* c.2753T>C	7.7	Y	*01/06/2016*	*TP53* c.814G>A	1.47	Y
06/12/2017	*RET* c.2671T>G	9.1	Y	*01/06/2016*	*RET* c.2671T>G	2.26	Y
01/06/2016	*RET* c.2753T>C	9.1	Y	*06/07/2016*	*RET* c.2753T>C	2.28	Y
16/11/2016	*RET* c.2753T>C	10.1	n*	*26/04/2017*	*RET* c.2753T>C	6.80	Y
01/06/2016	*NRAS* c.182A>G	18.2	Y	*06/12/2017*	*RET* c.2753T>C	8.12	Y
01/06/2016	*TP53* c.814G>A	19.8	Y	*01/06/2016*	*NRAS* c.182A>G	10.30	Y

Table 4-8. Detection of variant in plasma through ThyMa sequencing, ranked according to variant allele fraction (left side) and cell free DNA concentration (right side). *AF: variant allele frequency expressed as %. Det: was variant seen in plasma sequencing data, Y: yes, n:no (also highlighted in red). [DNA]: cell free DNA plasma concentration in ng/ µL. *: two outliers with higher AF where variant was not called with ThyMa sequencing, both of these samples had relatively low cfDNA concentration (0.32 and 0.43 ng/µL respectively) which may account for this.*

4.4.3.3.2 Detecting novel variants through plasma sequencing using ThyMa panel

Analysing the plasma sequencing data for new variants, eight potential candidates were present in more than one plasma sample from the same patient. Two databases were interrogated to see if these variants have been previously described, COSMIC and gnomAD. The results are summarised in **Table 4-9**. Results for patients with only one plasma time point are not shown in the table as it was felt that a variant appearing in only one sample would be less likely to represent a variant of interest, *i.e.* a driver mutation. Results were excluded if they occurred in homopolymer repeat regions, as sequencing error in these regions is an inherent weakness of the Ion Torrent sequencing technique.

Patient	Variant	Plasma 1	Plasma 2	Plasma 3	Plasma 4	Plasma 5	COSMIC	gnomAD
THY016	***KMT2C* c.2468T>C**	***nd***	**0.03**	**0.04**	***nd***	**0.03**	✓	X
THY015	*RET* c.2065T>A	*nd*	0.10	*nd*	0.09	x	X	X
THY015	***TSHR* c.775A>G**	***nd***	**0.05**	***nd***	**0.07**	**x**	✓	X
THY028	*KMT2C* c.2578C>T	0.41	0.36	x	x	x	✓	X
THY028	*KMT2C* c.2573G>T	0.40	0.36	x	x	x	✓	X
THY028	***KMT2C* c.2468T>C**	**0.04**	**0.04**	**x**	**x**	**x**	✓	X
THY002	*RET* c.2065T>A	0.12	0.03	x	x	x	X	X
THY044	*APC* c.2880delA	0.42	0.42	x	x	x	✓	X

Table 4-9. Novel variants detected in plasma. *TAF: tumour allele frequency. COSMIC: variant listed as known variant in database. gnomAD: variant is listed on genome Aggregation Database, representing a possible common SNP. nd: not detected in plasma. x: no plasma sample run. Brown colour highlight: candidate variants most likely to represent genuine mutations in ctDNA*

4.5 Discussion

In the below section the performance of the ThyMa panel with tumour samples relative to the GI2 panel is discussed, followed by the novel variants detected. This is then followed by a discussion of the panel's performance in plasma relative to dPCR results and then the novel variants detected in plasma. Finally, I discuss challenges associated with the project.

4.5.1 ThyMa panel: performance on tumour and comparison with GI2 panel

The primary aim of the validation experiment was to compare the performance of the targeted thyroid cancer specific ThyMa panel against the pre-existing colorectal GI2 panel. Both Gi2 and the ThyMa were shown to provide comprehensive coverage of the genomic aberrations in a significant portion (>80%) of thyroid cancer tumours. The new panel performed favourably when compared to the GI2 panel and detected variants in 9% more samples. This would have a significant positive impact on detection rates if used in future studies. Nine tumour samples were not run with the ThyMa panel due to either lack of sample or failure of LP QC. All variants in these samples that were called by the GI2 panel were in loci also covered by the ThyMa panel and so it is highly probable that these would have been detected as well.

Further advantages over the GI2 panel include lower costs and faster workflow. The library preparation and sequencing cost using the amplicon based ThyMa panel was approximately £120 per sample, whereas for the capture-based GI2 panel it was around £300 per sample. Time for library preparation was also significantly shorter: two days compared to four. In

future trials with larger number of samples this could represent substantial cost savings and faster turn-around time for results.

Moreover, as the design of the panel was based on a contemporaneous literature review, it is more thyroid cancer specific. The total coverage of the ThyMa panel was 20.47 kb, which was approximately eight time less than GI2 panel which was approximately 160.22 kb. Despite the significantly larger cover, the GI2 panel detected fewer variants and so part of this coverage would appear redundant. As the ThyMa panel has a more focussed coverage, it generates less extraneous sequencing data, rendering downstream data analysis more efficient.

4.5.2 Novel variants detected in tumour samples with the ThyMa panel

Novel variants not detected with the GI2 panel were seen in three genes using the ThyMa panel. These, and their dPCR validation tests are discussed in turn below.

4.5.2.1 TERTp variants and dPCR validation

One important mutation hotspot for thyroid cancer that was lacking in the GI2 colorectal panel was the *TERT* promoter region, hence its inclusion in the new panel. TERTp mutations are relatively common in DTC, particularly in follicular subtypes, where it is found in up to 16% of samples (132). In this study *TERT* variants were found in the tumour samples at more than double the expected rate. This is likely due to patient selection as the majority of patients in

this study had metastatic disease, and *TERT* variants are associated with more aggressive disease (section **1.1.5.2**)

The *TERT* c.1-124C>T dPCR assay functioned well and confirmed the NGS findings in the tumours for all cases when DNA was available. Unfortunately, no droplet separation was seen on dPCR testing in the *TERT* c.1-146C>T assay with wild type controls, indicating assay failure. I was therefore unable to validate *TERT* c.1-146C>T variant in the tumour sample of patient THY042. The promoter region of *TERT* contains many homopolymer repeats and has high GC content (circa 80%) making primer and probe design for the dPCR assay difficult.

4.5.2.2 KMT2C variant and dPCR validation

Although good droplet separation was seen in the *KMT2C c.1373delA* with WT controls (section **4.4.3.2.3.2**), no mutant was detected in the tumour sample for THY044 using dPCR. It is likely that this is a functional assay, but no positive controls were available to confirm this. As an alternative for positive controls, synthetic oligonucleotides could be designed containing the variant and used *in lieu* of human DNA as positive controls. Given the variant AF of 0.07 this result may, however, represent noise or a sequencing error.

4.5.2.3 CHEK2 variant and dPCR validation

The *CHEK2* c.591delA variant discovered in the THY047 tumour was confirmed on dPCR testing. This variant is not reported on the genome aggregation database, gnomAD[1], which would suggest it is not a common, non-pathogenic, single nucleotide polymorphism (SNP), nor is it listed on COSMIC. It is, however, listed as pathogenic on clinVar[2], on the basis of multiple published studies. As frameshift mutations have a higher probability of impacting protein function, it is plausible that this represents a driver mutation. The detected variant AF was 0.88 with sequencing, and 0.95 with dPCR. At such high allele frequencies, it is likely that this represents either loss of homozygosity or copy number amplification.

From the previous study in chapter 3, using the GI2 panel, a *PTEN* c.634+5G>C variant was found in the tumour of the above patient. This SNV affects a splice site and is listed on COSMIC with a FATHMM prediction score of pathogenicity of 0.99, hence it was called as a variant. However, this *PTEN* variant had not been detected the patient's five plasma samples with dPCR. Yet on repeat testing of those same plasma samples, the *CHEK2* c.591delA variant was detected. This confirmation of detected ctDNA increases the total number of patients with detected ctDNA in the previous study to 29 in 42 (69%). This demonstrates the increased value of using a more thyroid-specific panel.

[1] Genome aggregation database, containing pooled exome and whole genome sequencing data sets
[2] Public, freely available, online database that aggregates information about genomic variation and its relationship to human health with supporting evidence. See *www.ncbi.nlm.gov/clinvar/*

4.5.3 Tumour samples with no detected variants on either panels

Five tumour samples had no calls with either the GI2 or ThyMa panels, although three of these were not run with the ThyMa panel due to lack of sample or failed LP. There are several possible explanations for the lack of calls in these samples. Two of the samples had low DNA concentration and were extracted from lymph node tissue, rather than primary tumour, that had both low cellular content and tumour percentage. This could make detection of variants technically more challenging. It may have also been due to the inherent selectivity of a targeted gene panel: these samples may have had a genomic aberration not detectable by either panels, due to either non-covered loci or type of aberration (e.g. fusions or copy number variations). Alternatively, in some of these cancers, carcinogenesis may be due to epigenetic processes such as methylation, rather than genomic aberrations.

4.5.4 Performance of the ThyMa panel on plasma sequencing

The primary focus of the validation experiment was assessing the performance of ThyMa on sequencing tumour FFPE samples in comparison to the GI2 panel. The secondary, more exploratory aim, looked at direct ctDNA sequencing from plasma. The study protocol for the project described in the previous chapter relied on design of dPCR assays based on the molecular characterisation of tumours. Over time, however, different subclones may exert dominance due to selection pressure, particularly from targeted therapies, which may not have been identified at time of tumour resection. This is particularly relevant for thyroid cancer where the disease is relatively indolent with a natural history many years. Being able to detect *de novo* variants in the plasma without recourse to solid tissue biopsy for genetic material, which may not be feasible or safe, could be extremely advantageous clinically.

Circulating tumour DNA present in plasma samples is frequently at much lower concentrations, and significantly more fragmented, than the DNA extracted from solid tissue biopsy samples. Amplicon based TE, however, performs more favourably at lower concentrations compared with capture based, and so I attempted sequencing of the plasma samples.

As expected, plasma cell-free DNA concentration and ctDNA variant AF appear to affect the likelihood of detection with ThyMa sequencing. When samples had a concentration above 0.68 ng/µL all expected variants were detected. Variants with a TAF above 2.6% were generally detected, although this was less reliable than with overall concentration. The *RET* c.2753T>C variant in both plasma samples of THY048, with TAFs of 7.2% and 10.1%, were not detected. This may have been due to lower cfDNA concentrations in those samples or due to experimental error.

Given the relatively low number of samples, and broad ranges of concentrations and TAFs, this is only a coarse estimation of the limits of detection in plasma. If consideration is being given to using this panel for sequencing plasma in clinical studies, further experimentation should be performed to refine these estimations, using a higher number of samples, with concentration and variant AF ranges closer to the estimated levels.

4.5.5 Novel variants detected in plasma samples with the ThyMa panel

The panel detected multiple potential novel variants in the plasma. The gnomAD database lists known SNPs, and none of the variants shown in **Table 4-9** were listed on it, raising the possibility that these variants may be pathogenic.

Three of the variants, KMT2C c.2578C>T, KMT2C c.2573G>T and APC c.2880delA were found in the plasma at very high allele frequencies, 0.40 or above. Given these high AFs it is unlikely that these are ctDNA results and more likely represent detected germline variants in cell free DNA. This could be confirmed with buffy coat sequencing.

The variant *RET* c.2065T>A was seen in the plasma of multiple patients: two are shown in **Table 4-9**, and it was also seen in a further two patients. There is no published data on this variant as a disease-causing mutation in thyroid cancer. It would seem unlikely, therefore, to be a driver mutation in so many of this cohort and is more probably a sequencing error. The *KMT2C* c.2468T>C variant was also seen in multiple patients: THY016 and THY028. However, it is referenced in COSMIC as a pathogenic variant, and so this may represent a genuine novel variant.

The remaining variant was *TSHR* c.775A>G in THY015. The *TSHR* gene codes for a transmembrane receptor involved in transduction of TSH signalling to thyroid cells. Mutations in TSHR affect thyroid cell proliferation and function, and are associated with adverse outcomes in DTC when present in high allelic fractions (133, 134).

Thus, the most likely variants to represent potential novel subclones in plasma samples assayed were *KMT2C* c.2468T>C and *TSHR* c.775A>G. Two patients had more than two plasma samples sequenced (THY015 and THY016 in **Table 4-9**). Whilst the novel variants were not detected across all plasma time points, this may have been due to the variant AF dropping out of detectable range. The dPCR tracked ctDNA in both these patients in the previous study (Chapter 3) showed upwards progression in the allele fraction due to disease progression. This was not mirrored in the AFs of the *KMT2C* c.2468T>C and *TSHR* c.775A>G variants discovered here, perhaps as these new tumour variants do not represent emerging dominant subclones.

4.5.6 Challenges with a targeted panel

Both the GI2 and ThyMa panel cover loci involved in a significant portion of frequently arising mutations in thyroid cancer. Coverage of just five gene hotspots (*BRAF*, *RET*, *NRAS*, *HRAS* and *KRAS*), would have led to detection of mutations in 35 of 51 patients (68%) in this cohort. Whilst aberrations in these genes are frequently responsible for driver mutations in thyroid cancer, recurring mutations in other genes are significantly less common. **Figure 4-9** illustrates the 20 most frequently involved genes in thyroid cancer: there is an almost exponential reduction in frequency of involved genes. This makes design of a targeted gene panel, which is inherently selective by design, problematic: a balance of increased detection rates against increasing spurious loci coverage and costs must be maintained. However, in the future, as costs of sequencing and efficiencies in data handling continue to improve, this issue will become less relevant.

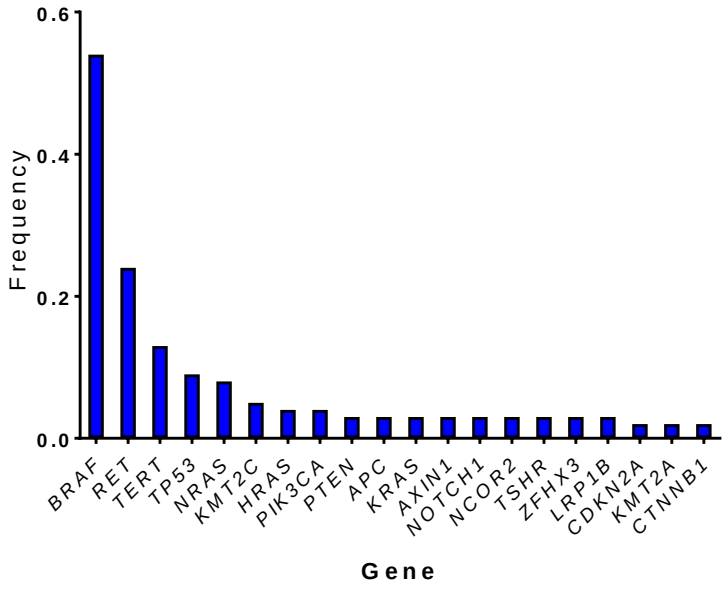

Figure 4-9. Top 20 genes with pathologic variants detected in thyroid cancer across all histological subtypes*. Listed in descending order of frequency. Data from COSMIC database*

The ThyMa panel focusses primarily on one type of genomic aberration, namely SNVs and InDels. Other types, such as copy number variations (CNVs) including amplifications/deletions, gene fusions and epigenetic changes such as methylation are not covered. A targeted gene panel, by definition, does not cover all possible loci or types of aberrations and by necessity of cost and efficiency is selective. Whilst it was felt that this panel design offered good return in terms of coverage, clearly some of the aforementioned types of aberrations are causes of genetic change in thyroid cancer (20, 28) and are not covered by the panel. They may account for a portion of driver mutations in the tumour samples in which no variant was detected using the ThyMa panel.

Whilst technically possible, determining CNV of genes based on data from a panel with a relatively small coverage is unreliable and thus was not attempted for this study. A much

larger panel, one that includes multiple reference levels from stable regions such as those in close proximity to centromeres, or whole exome sequencing (WES) would better allow for this but would substantially increase costs, data output and material required.

Fusion breakpoints are often located in non-coding areas, e.g. at intronic or intergenic locations. Breakpoint location within introns renders gene-fusion detection with amplicon-based TE problematic. Without prior knowledge of the fusion breakpoint to construct appropriate amplicons around this locus, the entire intronic or intergenic region would need to be tiled with amplicon coverage. This would significantly increase the cost of running the panel. This issue can be partly overcome with capture-based TE: if the baits align on one side of the breakpoint, then the fusion sequence may be captured. For more consistent results, however, the use of RNA-sequencing would obviate the requirement for the breakpoint, as would Whole Genome Sequencing (WGS) but RNA is inherently much less stable than DNA and can be difficult to extract from FFPE and plasma, and WGS is currently vastly more expensive.

4.5.7 Limitation of plasma sequencing experiment

When compared to tumour samples, allele fractions of tumour variants in plasma are often very low (0.01%) and the DNA concentration in the sample is usually small. This makes variant calling difficult: e.g. if a variant AF is only 1% (*i.e.* only 1 in 100 of the pre-PCR DNA fragments at that locus contain the variant), and the sequencing depth is 10x then stochastically it is easy

for the variant to be missed. Increasing the number of PCR cycles to allow a greater read depth helps circumvent this problem.

To deal with allele fractions as low as 0.01% in plasma, ultradeep sequencing of up to 10,000x has been advocated, however, this increased number of PCR cycles required for this raises the risk of PCR duplicates. This makes it harder to ascertain if variant calls and changes in AF are genuine. Ideally, all DNA fragments are amplified equally during the PCR process yet unequal amplification can occur of a DNA fragment, leading to PCR duplicates and giving the illusion that a DNA fragment accounted for a significantly higher proportion of reads at the start than it did (so the AF might appear as 3% instead of 0.01%). Thus, the new variants detected in the plasma in this experiment (**Table 4-9**) should be confirmed with dPCR to ensure they are not the result of PCR duplicates. This was not feasible here due to time constraints.

A new sequencing technique using molecular barcodes could obviate the need for dPCR validation in future studies and allow for more confident calling in plasma. In this method, unique molecular identifiers (UMIs) composed of individual nucleotide sequences, are ligated to each DNA fragment in a sample prior to PCR amplification. Reads that have occurred due to PCR duplicates of a single fragment would all have the same UMI and can thus be recognised and accounted for by the bioinformatics pipeline.

Greater sequencing depth and the use of UMIs would allow for more reliable detection of variants with low AF in plasma samples. If possible, this should be used in future studies

looking to sequence from plasma, but this is a relatively novel technique and the costs involved are greater.

Chapter 5. Mechanisms of resistance to targeted therapies in advanced thyroid cancer

5.1 Introduction

As described in the introductory chapter, several multi-kinase therapies have received regulatory approval for use in advanced thyroid cancer. These drugs do not kill cancer cells, they are not cytotoxic and cannot cure, but through blockade of essential pathways can render the cells non-proliferative and prevent progression of disease. They are thus deemed to be cytostatic. Recognition of this mechanism of action is important as it requires that the therapy be continued indefinitely, until limited by either the development of resistance or intolerable side effects. Four multi-kinase inhibitors have been approved by the FDA and the EMEA for thyroid cancer: lenvatinib and sorafenib for radioiodine refractory DTC, vandetanib and cabozantinib for MTC. In the UK, NICE has recommended the use of three of these. However, in guidelines published in December 2018, it did not recommend vandetanib due to perceived unclear clinical benefits (135).

Whilst these targeted therapies can be effective in the short term and increase progression-free survival, this effect is not durable. This a common theme for this class of drugs across cancer types: most tumours will develop secondary resistance to these drugs within 12 months, sometimes even within weeks (136). Furthermore, a significant percentage of patients will either fail to respond or respond poorly due to primary resistance. Lenvatinib, for example, increases progression free survival by 14 months over placebo with a response

rate of 65% (137). As a result, a re-alignment in expectations of how these drugs perform, particularly as monotherapies, has necessarily occurred.

Resistance to these therapies is a challenge to achieving durable response. Yet, there is a paucity of data concerning the mechanism(s) leading to resistance to these drugs in thyroid cancer, whether these be genomic, epigenetic or other aberrations. Given the nature of cancer, with rapid cellular division and unstable genomes prone to changes in key cell-cycle regulatory genes, it is likely that drug resistance due to selection pressure is a predictable event. As a result of this evolution of tumour cells, trialling individual drugs in sequence may actually be detrimental and promote clonal expansion of existing resistant cells. Better knowledge of resistance mechanisms may allow for selection of more effective therapies, particularly if combined with contemporaneous molecular characterisation of tumours. It would also allow for prompt cessation of a futile treatment and switch to an alternative effective therapy if the resistance mechanism becomes detectable in a treated patient.

Two types of resistance to targeted therapies are recognised. Primary resistance is present within the tumour cells prior to treatment commencement. Secondary resistance is that which is acquired during treatment. Some data do exist with regards primary resistance in thyroid cancer. Vemurafinib is a selective *BRAF* inhibitor that has been effective in achieving a response in over 50% of malignant melanoma tumours harbouring the *BRAF* V600E variant (138). However, its response rate in mutant *BRAF* PTC was unexpectedly lower, with only 35% achieving a partial response (139). Several mechanisms have been postulated to account for this primary resistance to *BRAF* blockade, with mechanistic data to support them. These

include copy number gain in Myeloid Cell Leukaemia 1 (*MCL1*) gene, an anti-apoptotic protein (140) and increased HER3 signalling leading to rebound increase in downstream elements of the MAPK pathway (141). As touched on previously, *BRAF* V600E mutation also diminishes iodide uptake by follicular cells through upregulation of epigenetic pathways that reduce expression of the sodium/iodide symporter, possibly a cause of primary resistance to RAI therapy (142). In MTC, resistance to cabozantinib has been associated *in vitro* with *RET* V804M, a variant which prevents the drug from binding its target receptor (143).

Secondary resistance is less well characterised in thyroid cancers. One study has looked at a PTC cell line that developed resistance *in vitro* to pazopanib, a multi-kinase inhibitor, and found a *KRAS* G13V variant (144). This variant is known to be associated with constitutive KRAS activation (145) which could represent a putative resistance mechanism. However, this initial finding has not been supported with mechanistic or *in vivo* studies and, further, pazopanib is not used routinely for treatment of thyroid cancers. As yet, there are no known published causes of secondary resistance *in vivo* to targeted therapies in thyroid cancer. In contrast, several genetic causes of resistance to inhibitors are well documented in other malignancies. As mentioned previously, *ESR1* mutations are known to confer resistance to aromatase inhibitors in breast cancer, and the T790M *EGFR* variant is an established cause of targeted therapy resistance in NSCLC. Both have been shown to be detectable in ctDNA and may be used to guide treatment selection. A clearer understanding of the resistance mechanisms in thyroid cancer would clearly also be hugely beneficial, especially if detectable in plasma.

It is within this context that this study was conducted, and secondary resistance was the focus. Sequencing of DNA samples taken pre- and post-resistance is an experimental technique that can be used for discovery of genetic causes of secondary resistance. This approach was used effectively in chronic myeloid leukaemia to discover resistance-conferring mechanisms to imatinib (146) and allowed for the development of alternative targeted therapies. Given the lack of prior data relating to secondary resistance in thyroid cancer, this necessitated an agnostic approach to increase the chances of detection. Much broader genome coverage is required, therefore Whole Exome Sequencing (WES), rather than a targeted panel, is used in this experiment. Using WES also allows for copy number variations to be detected, which is more difficult with targeted gene panels.

In this study, samples are used from patients known to have responded to targeted therapies and subsequently progressed due to resistance. In sequencing samples taken prior to starting treatment and after resistance has developed, it was hoped that novel genomic aberrations appearing in the latter sample may represent putative secondary resistance-causing candidates. The DNA was obtained from samples remaining used in the ctDNA study described in chapter 3. This experiment was planned as a pilot study to be used as a proof-of-principle for a larger follow-on study if the workflow was shown to be feasible.

5.2 Hypothesis

Secondary resistance to targeted therapies in thyroid cancer occurs due to acquired genomic aberrations, and these may be detectable in circulating tumour DNA.

5.3 Aims

To discover novel genomic aberrations that are candidate mutations responsible for secondary resistance to targeted therapy in thyroid cancer through WES sequencing of paired DNA samples taken prior to starting therapy and after resistance has developed.

5.4 Results

From the cohort of patients entered into the study in chapter 3, four patients were identified that had paired samples in which ctDNA was detected after a period of being on a TT that either halted progression or led to a partial response, as seen on axial imaging (SD or PR by RECIST criteria) followed by disease progression (PD). Three of these were paired tumour-plasma samples and one was plasma-plasma. Thus, eight samples were sequenced and the details of these are summarised in **Table 5-1**.

Trial ID	Type	Sample ID	Date	Type	TT	TT Dates
THY041	*MTC*	16/07997 16/08315	2001 06/2016	*Tumour* *Plasma*	Vandetanib	01/2015 – 12/2015
THY016	*MTC*	16/08311 18/01288	07/2016 12/2017	*Plasma* *Plasma*	Vandetanib	08/2016 – 10/2017
THY002	*MTC*	17/07001 18/02010	2014 10/2017	*Tumour* *Plasma*	Vandetanib	09/2015 – 07/2017
THY015	*MTC*	17/07007 17/11262	2005 09/2017	*Tumour* *Plasma*	Vandetanib Nintedanib Cabozantinib	03/2013 – 12/2016 01/2017 – 03/2017 05/2017 – 11/2017

Table 5-1. List of samples selected for WES, grouped per patient*. Date: date tumour/plasma sample taken. MTC: Medullary Thyroid Carcinoma. TT: targeted therapy. TT dates: dates patient on targeted therapy*

The sequencing run generated a total of 756,483,151 reads. The mean depth across all samples was 88x. The median number of on target reads was 70,145,356 (70.3%) per sample.

5.4.1 Post-resistance changes: SNVs & InDels

The first part of the analysis focussed on SNV & InDels that were detected in more than one patient. These variants are shown in **Table 5-2**, with allelic fraction when detected. The first round of analysis looked at variants that were detected in all 4 patients only, of which three were found, and these are the first three patients listed in **Table 5-2**. The search criteria was then widened to include variants that were found in two or three patients, which hugely increased the number of variants calls for review. Therefore, to allow for manual review of each variant I was more selective with the filtering and excluded all calls in non-exonic regions. The evidence supporting the candidacy of the remaining variants as resistance-conferring mutations is described in the discussion section.

Gene	Variant	Function	*Sample variant AF*							
			THY041		***THY015***		***THY002***		***THY016***	
			Pre	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
ACTR3BP5	Complex del	*pseudo*	*x*	0.09	*x*	0.10	*x*	0.13	0.20	*x*
MUC16	GA>TG	*intronic*	*x*	0.13	0.08	0.13	*x*	0.54	0.06	0.10
CDC27	c.77T>C	*missense*	0.72	x	0.82	0.19	0.79	0.10	0.13	X
GPR1	c.409A>C	*missense*	*x*	0.06	*x*	*x*	*x*	0.16	*x*	*0.10*
HYDIN	c.7588A>G	*missense*	*x*	0.17	*x*	*x*	*x*	*x*	0.23	0.21
MRI1	c.893A>C	*missense*	*x*	0.07	*x*	*x*	*x*	*x*	*x*	0.05
CDC27	c.80T>C	*missense*	*0.72*	*x*	0.81	0.19	0.80	0.10	*x*	*x*
MUC4	c.6746C>T	*missense*	*x*	0.12	*x*	*x*	*x*	*x*	0.16	0.15
FDFT1	c.193_198del	*deletion*	*x*	*x*	1.00	1.00	0.99	1.00	*x*	*x*
KIR3DL3	c.90C>G	*missense*	*x*	*x*	*x*	*x*	*x*	0.03	*x*	0.04

Table 5-2. List of variants detected in at least two patients*. Variant: description of mutation. Function: location or effect of variant. Pseudo: pseudogene. Pre: variant allele fraction (VAF) in first (pre-treatment) sample. Post: VAF in second (post-resistance) sample. x: variant not detected*

The first variant is a 36 bp deletion seen in Actin-Related Protein 3 Homolog B Pseudogene 5 *(ACTR3BP5)*. It is seen in the post-resistance sample only in three of the four patients, however, ACTR3BP5 is a pseudogene. Pseudogenes are non-coding sequences of DNA that resemble a gene but are not translated into functional protein. They are probably the result of previously active genes that have mutated to inactive form. Functionally, therefore, this variant is an unlikely candidate for resistance. The disappearance of this variant in the post-resistance sample for patient THY016 also detracts from the probability of this. The presence of a variant with such a large deletion in multiple samples may be explained by read alignment error.

The second variant is an InDel in Mucin-16 *(MUC16).* In normal physiology, *MUC16* encodes a transmembrane protein that is a member of the mucin family involved in providing a protective mucous barrier in epithelial cells. However, it is a known tumour marker in ovarian cancer and was previously known as Ovarian Carcinoma Antigen or CA125.

The third variant is a missense mutation seen in Cell Division Cycle 27 (*CDC27*): c.77T>C. The *CDC27* protein is a core subunit necessary for the activation of the Anaphase-promoting complex/cyclosome (APC/C). APC/C interacts with mitotic checkpoint proteins and is involved in cell-cycle regulation.

On review of current literature, *KIR3DL3*, *FDFT1* and *HYDIN* genes have very little supportive data published in favour of playing a significant role in cancer cell biology. They are not known to be involved in downstream elements of blocked pathways or even more broadly in cell

cycle control, and so have a low likelihood of being involved in potential mechanisms for resistance to targeted therapies. Thus the remaining pertinent variants from **Table 5-2** are *CDC27* c80T>C, *MUC4* c.6746C>T, *MRI1* c.893A>C and *GPR1* c.409A>C. and these are discussed later.

5.1.1 Post-resistance changes: Copy Number Variation

Copy number variation was assessed across the samples, and a heatmap summarising the results is shown in **Figure 5-1**. No significant changes in were seen in post-resistance samples.

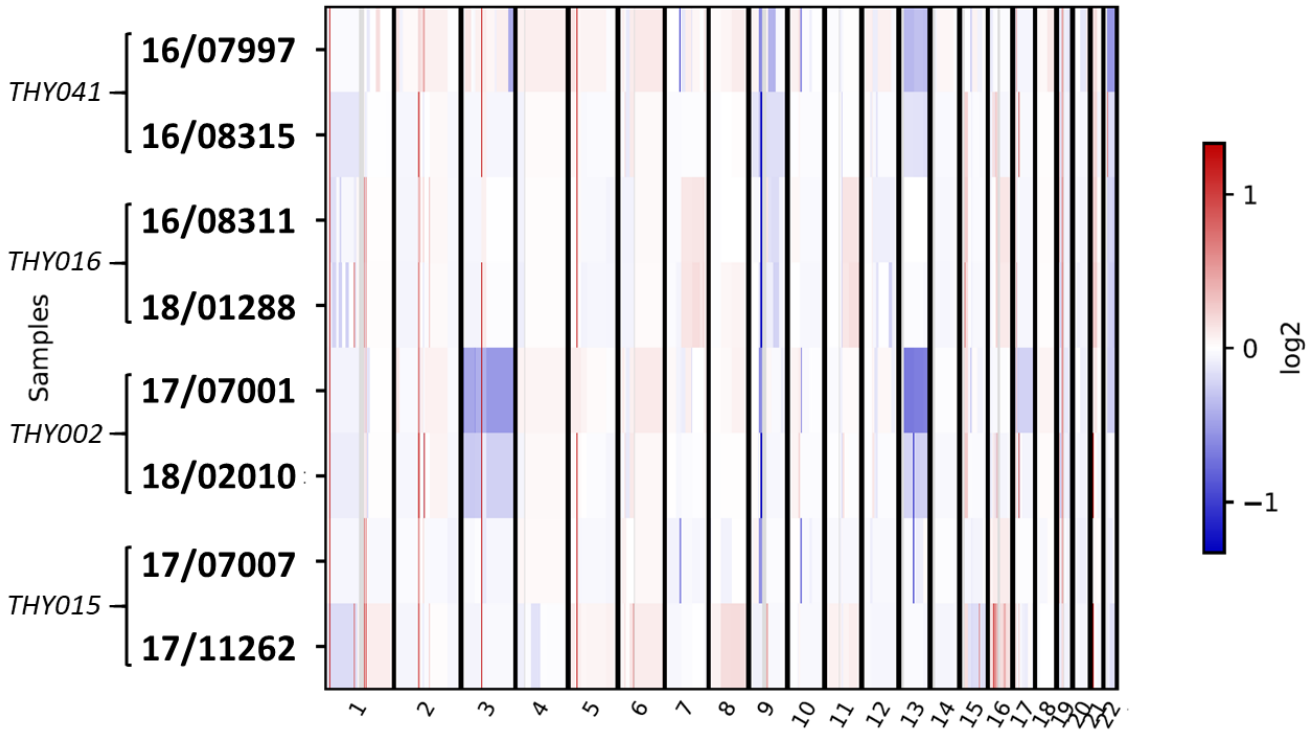

Figure 5-1. Heatmap of copy ratio (log2) for all chromosomes per sample

5.5 Discussion

This experiment was a proof-of-concept study that fulfilled its aim to establish if the methods used could offer a viable workflow for discovering genomic aberrations associated with secondary resistance to TT in thyroid cancer. It has also usefully highlighted areas for improvement that, if addressed, would allow for better quality future studies. Furthermore, it has identified interesting leads in terms of putative resistance-conferring variants that would be ideal for further study and these are discussed below in order of decreasing plausibility.

5.5.1 Putative resistance conferring candidates

This study selected patients with known resistance to targeted therapy, and several variants were identified that occurred across multiple patients that were in either the post-resistance sample only or in both samples. Some of these were discounted as plausible resistance-conferring mechanisms in view of published literature concerning their known function. However, several variants were detected in genes that have been shown to play an important role in cancer and may be important in conferring treatment-resistance.

GPR1 c.409A>C

Chemerin is a multifunctional cytokine involved in adipogenesis and glucose homeostasis. Its function is primarily mediated via two cellular receptors, one of which is G-Protein Coupled Receptor 1 (*GPR1*). The role of chemerin in cancer is not yet fully understood and it appears

to have both tumour-suppressive and cancer-promoter effects, the balance of which varies between cancer types (147).

Its role in cancer appears to be mediated primarily via several mechanisms. Firstly, it acts as a chemo-attractant for anti-tumour effector immune cells such as NK cells that are key in the immune response to cancer (148). Secondly, it is involved in intracellular signalling and there is evidence that its tumour suppression activity is related to activation of PTEN (149) and inhibition of the MAPK pathway (150). Conversely, it can activate the MAPK pathway leading to tumour promotion in gastric cancer (151). Lastly, chemerin also has potent angiogenic effect *in vitro* on endothelial cells and so may promote tumour angiogenesis (152)

Expression of chemerin or its receptors have been found to be altered in several cancer types. In breast cancer, high mRNA expression in tumour cells of GPR1 receptor was found to be strongly associated with a longer disease-free survival (147). In NSCLC patients with lower chemerin expression had poorer survival rates (153). In contrast to this however, chemerin appears to have a tumour-promoting effect in gastric cancer, where increased serum levels are associated with advanced clinical stages (151) and shorter overall survival (154).

In relation to thyroid cancer there is very little published data. One study has found that 5-year survival rates of thyroid cancer patients was significantly lower in those with higher serum chemerin levels (155).

As GPR1 is an important receptor for chemerin, a missense mutation in *GPR1* could lead to altered downstream activation of effector pathways. From what is known about its role in cancer it is plausible that this could impact on TT resistance. The *GPR1* c.409A>C variant was found de novo in three of the patients in the later sample, and as such would warrant further investigation.

***MRI1* c.893A>C**

One mutation seen in Methylthioribose-1-Phosphate Isomerase 1 (*MRI1*) c.893A>C is present as a novel variant in the second sample in two patients. *MRI1* codes for an enzyme involved in amino acid synthesis that is used in the methionine salvage pathway. There is very little data published on the role of MRI1 mutations in cancer. However, increased levels of expression are associated with metastatic melanoma, and it appears to promote cell invasion through activation of Ras Homolog Family Member A (RHOA). *RHOA* is involved in actin cytoskeleton organisation and regulates cell shape and motility, overexpression is associated with tumour metastasis (156). This variant is seen in the post-resistance sample only for two patients and therefore would also be an interesting candidate to pursue.

MUC16 c.GA>TG

Increased expression levels of MUC16 have been observed in many cancer types, including breast, bladder and pancreatic, and is associated with a worse prognosis (157-160). It has a variety of cancer-promoting effects, such as encouraging metastasis through alteration of cell-cell adhesion and migration (161), as well as immune-modulation by inhibiting Natural

Killer (NK) cell activity against cancer cells (162). Of particular interest, however, is that changes in *MUC16* expression have been shown to activate EGFR *in vitro* and lead to increased phosphorylation of the downstream elements Akt and ERK1/2 (163). This is a pathway targeted by the multi-kinase inhibitors used in thyroid cancers. Further, MUC16 changes have been shown to contribute to resistance *in vitro* to cytotoxic drugs in pancreatic and ovarian cancer cell lines (158, 164). Biologically then, mutations in this gene could plausibly contribute to resistance in thyroid cancer.

This variant was seen *de novo* in two of patients in the post-resistance samples. In the other two patients it is seen in both samples but does show an increasing trend in the AF. This may be represent a pre-existing variant that is being selected for due to sub-clonal evolution as a result of conferring resistance. Thus, this could be a potential candidate from a timing perspective. Whilst the *MUC16* variant detected here was in a non-coding region, intronic changes can result in altered gene expression levels if located in regulatory regions through introduction of novel splice sites or activation of novel promoters.

MUC4 c.6746C>T

Mucin 4 *(MUC4)* is also a member of the mucin family mentioned previously, and codes for a protein involved in the formation of a mucus barrier for some epithelial cell types. However, it may play a role in tumour progression as a result of its effect on cell proliferation and, similarly to *MUC16*, is also over-expressed in some cancers. *MUC4* acts as a ligand and activator for the receptor tyrosine kinase *ErbB2*, also known as Human Epidermal Growth

Factor Receptor 2 (*HER2*). It has been shown that an activated *MUC4-ERBB2* complex leads to down-regulation of Cyclin Dependent Kinase Inhibitor 1B (*CDKN1B*) which represses apoptosis and stimulates cellular proliferation (165, 166). This could be a contributory mechanism to developing resistance. In line with this, *MUC4* has been found to be over-expressed in breast cancer cell lines resistant to trastuzumab (167) and *MUC4*-deficient pancreatic adenocarcinoma cell lines were shown to be more sensitive to gemcitabine in vitro (168). Given this data in other cancer types relating to therapy-resistance, including TT, it is possible that a similar effect could be occurring in thyroid cancer. In this study, a MUC4 c.6746C>T variant is seen *de novo* in the second sample of THY041. In THY016 it is seen in both samples, but as the time span between samples was significantly shorter, this may be a variant that emerged earlier, prior to starting TT.

CDC27 c.77T>C and CDC27 c.80T>C

As discussed previously the *CDC27* protein is a subunit of the Anaphase-promoting complex/cyclosome (APC/C) that interacts with mitotic checkpoint proteins and is involved in cell-cycle regulation. Mutations in *CDC27* have been detected in various cancers, in particular colorectal. It is frequently upregulated in cell lines and tumours and is associated with increased TNM stage, distant metastasis and decreased survival (169). Interestingly, dysfunction of APC/C activity confers resistance to kinase inhibitor therapy used in breast cancer, possibly by minimising the lethal effects of mitotic segregation errors (170). Therefore, although there are no published data currently, it would seem plausible that a mutation in *CDC27*, as a key part of APC/C, could lead to APC/C dysfunction and in turn resistance to targeted therapies. The *CDC27* c.77T>C variant is listed on the COSMIC database,

where it has an association with thyroid malignancy. Its mutational impact is calculated as 0.95 (high) using FATHMM-MKL, an algorithm that predicts the functional and phenotypic consequences of missense variants using computer based modelling (171). The AF of the c.77T>C variant in the samples appears to drop in the later samples for THY015 and THY002, or is not present (THY016 and THY41) which suggests it is unlikely to represent a variant leading to secondary resistance. However, conclusions based on allele frequency with differences in tumour purity between samples is difficult, a point expanded on later. It may contribute to primary resistance.

5.5.2 Highlighted learning points from study

5.5.2.1 Inter-sample timespan

The time elapsed between earlier and later samples in the study was varied across the four patients. The range was from 17 months to 15 years. Whilst thyroid cancer, particularly DTC, is often relatively indolent and can take years to progress, it would be difficult with the longer spans to assume that novel mutations had arisen during the TT period only. A 15-year period is likely to be enough time for multiple new variants to occur spontaneously due to evolution of the cancer. Where possible, paired plasma sequencing was performed rather than tumour-plasma sequencing, due to the shorter time span between samples. This was only possible with one patient, although in one other patient the tumour sample used was taken only one year prior to starting TT. Furthermore, one of the patients had been on multiple TTs during the time frame, which makes it difficult to attribute resistance to a single agent. A plasma sample taken concurrently with the initiation of each therapy would represent the ideal “pre-resistance” sample. A further sampling on detection of disease progression (DP) on axial

imaging should also be taken as the "post-resistance" sample. This reactive approach to sampling should be performed in future studies, rather than at pre-determined time span.

5.5.2.2 Sample size

Four patients from the remaining samples from the study in chapter 3 fulfilled the inclusion criteria for this study. Whist not large numbers, in other cancer types single variants can account for a large proportion of resistance-mechanisms seen. For example, the T790M *EGFR* mutation in NSCLC is seen in 60% of cases of resistance to EGFR blockade (172). It was therefore felt that proceeding with this cohort would be interesting to generate possible hypotheses for future studies. Clearly a larger number of samples is required to increase the probability of detecting a candidate variant.

5.5.2.3 Sequencing depth

The mean sequencing depth for the samples in this experiment was 88x. Sequencing from plasma at this relatively low depth makes variant calling less certain, as it becomes challenging to distinguish rare variants from sequencing errors. To illustrate this, I manually reviewed the sequencing data from the plasma samples in this experiment to determine if the ctDNA variants for those samples identified in chapters 8 and 9 were seen. All variants were seen in the BAM files but were not identified by the *Mutect2*[3] caller used in the WES pipeline. Thus, the variants were present in the data but lacked sufficient reads (often 10 or under) for the caller to list them as positive calls. Sequencing to a much higher depth and using UMIs would

[3] Mutect2 is a publicly available somatic variant caller published by the Broad institute as part of the genome analysis toolkit, frequently used as part of the pipeline used to analyse sequencing data.

significantly improve this and allow for more confident calling, and should be used in future studies. This would, however, increase the cost per sample considerably.

5.5.2.4 Tumour purity

Tumour purity (TP) refers to the proportion of cancer cells (or DNA if related to plasma) in a sample relative to normal cells/DNA. This is important to note if comparing variant allele fractions (VAF) between samples with different purities and makes it more difficult to draw conclusions about dynamic changes in VAFs. Clearly there can be a substantial difference in TP between tumour samples, where it may be 10-90%, compared with plasma samples, where ctDNA is often only a small fraction of total cell-free DNA. In this study, whilst tumour-plasma comparisons have been made, I have drawn no significant conclusion based on VAF changes. In future studies, to allow for AF comparisons over time, plasma-plasma sampling would be preferential. Differentials in TP could be controlled for using bioinformatic techniques, although this remains an evolving field. How best to account for variations in the volume of ctDNA shedding by the tumour, as the disease changes, still needs to be addressed.

5.5.2.5 Buffy coat usage

Finally, buffy coat samples were not used as matched controls for these patients. Previously obtained WES data from a patient in a different study was used as reference genome instead to reduce cost. This does make CNV estimation less accurate, as the samples would be normalised against to a reference genome that is not their own. Use of patients own buffy

coats also helps to identify variants that are germline SNPs. Ideally, patients should have a third control sample taken from the buffy coat to ensure more accurate results.

5.5.3 Validation of findings and broader view

5.5.3.1 Confirming variant as resistance-conferring

The variants identified in this small samples require further data to support the hypothesis that they confer treatment resistance. This could be done using novel techniques such as the CRISPR/Cas9 gene editing tools. The CRISPR/Cas9 system is a part of an adaptive immune response to invading viruses found in prokaryotic cells, and it has been harnessed to allow genome editing in mammalian cells. The CRISPR-associated protein 9 (Cas9) enzyme is a nuclease able to recognise and cleave specific sequences of DNA. Cas9 is guided to the location of interest by a 20 bp guideRNA (gRNA), and specific gRNA sequences can be created *in vitro.* Once cleaved, the DNA is repaired by normal cellular machinery and customised donor DNA sequences can be introduced into the genome, thus allowing specific mutations to be inserted (173). This technique has been further refined, and Cas9 variants have been engineered that can induce single nucleotide changes, and have been used to construct cancer cell lines with specific mutations (174). This system has already been exploited to uncover novel mechanisms of resistance to targeted therapies such as imatinib (175) and bortezomib (176) in CML cell lines. This system could be used to create a thyroid cancer cell lineage with the identified putative variant. Testing of these cells in vitro could then be conducted to ascertain therapeutic resistance: positive results would offer supportive data that the candidate aberration was a plausible mechanism.

5.5.3.2 Detecting other mechanisms of resistance

A large proportion of known causes of resistance to therapies, such as those described in the introduction, are activating exonic SNPs. For example, the *EGFR* T790M mutation is seen in up to 50% of lung cancer resistance to EGFR blockade. Copy number changes have also been found to be contributory to resistance, such as with *MET*, *HER2* and *EGFR* amplifications. Both SNPs and CNVs are detectable with whole exome sequencing, hence this method was chosen.

However, whilst WES is relatively agnostic compared to targeted gene panels such as ThyMa, it does not offer good coverage of non-exonic areas. Variants such as splice junctions, regulatory sequences (e.g. promoter or silencer regions) and intergenic regions are not covered. This cohort of aberrations may therefore not be detected, although as yet there is little data to suggest that variants in these loci lead to resistance. Conversely, gene fusions are known to contribute to therapy resistance: *ALK* fusion products can lead to resistance to targeted therapy in lung cancer. Identification of gene fusions is also sub-optimal in WES, however to cover the majority of genomic aberrations would require whole genome sequencing (WGS) and this would increase both costs and data output. An alternative, that can be used in tumours, is sequencing from RNA. However, RNA is an inherently more unstable molecule than DNA and RNA-seq from plasma is difficult with current techniques and poorly established as yet.

Another consideration is that the mechanism of resistance in thyroid cancers may not be driven by genomic changes. Epigenetic changes, such as DNA methylation and histone modification, can alter gene expression and these may also be involved in resistance. Sharma

et al. has shown that transient epigenetic changes led to TT tolerance in EGFR-inhibitor resistant NSCLC cell lines through upregulation of a histone demethylase (177). Given the lack of change to genomic DNA, Epigenetic modifications are not detectable with routine sequencing techniques. Detection of these requires modified assay techniques, such as bisulphite sequencing.

5.5.3.3 Primary resistance

This study focussed on discovery of secondary resistance mechanisms and the approach utilised, comparing pre-therapy and post-resistance samples, is likely a useful technique for this. However, this approach is less likely to be fruitful in searching for *primary* resistance mechanisms. As discussed previously, a proportion of patients are non-responders, or very poor responders, to targeted therapy. This would imply that a mechanism of resistance, genomic or otherwise, is already present in at least a significant portion of tumour cells in these patients. For primary resistance, unlike in this study, focusing on variants that are only present in the second sample would be erroneous as this would miss variants present in the first sample but selected for during targeted therapy that become dominant as a result of their phenotypic advantage. Primary resistance mechanism discovery would require a modified study design. For example, comparing pooled sequencing results from cohorts of responders to non-responders might allow common mutations in the non-responder groups to be identified.

Chapter 6. Conclusions and future directions

Assaying circulating tumour DNA may have several advantages over conventional tumour markers and traditional tumour genotyping with solid tissue biopsies. It is minimally invasive, contemporaneous and may capture tumour heterogeneity. This thesis set out to further the understanding of circulating tumour DNA in thyroid cancer. Broadly, the aims were to investigate the detection of ctDNA in patients with advanced disease to generate supportive data and identify exploratory avenues for future studies, as well as establishing a useful assay to support such investigations. In this regard the aims of the thesis were all met.

6.1 ctDNA in advanced thyroid cancer

This study was a multi-mutational analysis of ctDNA across a range of thyroid cancer subtypes over multiple time points. It demonstrated that, using a multi-variant approach, a high proportion of patients with advanced thyroid cancer have detectable levels of ctDNA in their plasma. A correlation of ctDNA with conventional markers of disease in DTC was observed and ctDNA concentrations were reflective of disease burden and disease progression. Although not statistically significant, a trend was observed through modelling where increases in ctDNA predicted disease progression on axial imaging, and it is likely that with a larger sample size this would have been significant.

The study also highlighted areas in which ctDNA may demonstrate superiority over currently used tumour markers. Based on the data, ctDNA may be particularly useful when

conventional markers are not helpful, for example in de-differentiated disease or when anti-Tg antibodies render Tg an unreliable tumour marker. ctDNA is, therefore, likely to complement current conventional biomarkers in the management of thyroid cancer. Further, the data suggest that ctDNA may predict disease progression earlier and more confidently than conventional markers, particularly in MTC. For patients on targeted therapies, ctDNA may offer a more rapid and reliable assessment of disease response, providing information on effectiveness of therapies and guiding drug selection.

These early results are promising and ctDNA may prove to be a clinically useful biomarker in thyroid cancer. Further investigation with a larger sample size is required to confirm these initial findings. The data generated through this study has been presented at the NCRI Head & Neck cancer subgroup meeting and a future multi-centre study was greenlighted. This is currently in development.

6.2 ThyMa Panel

For use in further ctDNA studies a thyroid cancer specific targeted gene panel would be advantageous, and so the primary aim of this study was to design such a panel. The ThyMa panel performed well in this validation study. The detection rate is 9% higher in tumour FFPE samples compared to the colorectal-focussed GI2 panel. Detection of new variants with this panel in tumours also led to the detection of ctDNA in a patient in which ctDNA had not been detected in the study described in the previous chapter. The panel has a simpler library

preparation workflow and is cheaper per sample than the GI2 panel. Drawbacks include lack of coverage for aberrations other than SNVs and InDels, such as CNVs and gene fusions.

Early results indicate that it could potentially be used for sequencing from plasma. Sequencing from plasma is beneficial because it allows for exploration of the complete mutational landscape of a tumour and does not require *a priori* knowledge of a variant to design an assay, unlike dPCR. However, sensitivity and accuracy at the very low concentrations and allele frequencies seen in plasma ctDNA mean that it is less accurate, and the very low allele fractions seen in ctDNA positive results in this study would need confirmation with alternative genotyping techniques, such as dPCR. A more attractive option for future studies would be ultra-deep sequencing to a higher depth (such as 20,000x) and making use of unique molecular identifiers (UMIs) pre-amplification to reduce PCR duplication error, although this is more costly. Given the positive performance of the ThyMa panel, it will be used for planned future studies by our unit.

6.3 Mechanisms of resistance to targeted therapies in thyroid cancer

This experiment demonstrated that whole exome sequencing of DNA samples taken pre- and post-resistance to targeted therapy may be a useful technique for identifying genomic mechanisms of resistance. Potential candidates worthy of further investigation were discovered, and this should be performed with methods as outlined in the discussion. The study also highlighted areas in the workflow that could be improved to allow for more successful future studies. A larger number of patients would increase the chances of detecting

a disease-conferring mutation. Paired tumour-plasma samples should be avoided, and plasma-plasma samples are preferred. Ideally these should be planned prospectively to be taken concurrently with starting treatment and when resistance is noted to a monotherapy. This would reduce time-span between samples, avoid confusion over which therapy is responsible, and reduce tumour purity comparison concerns. To allow for confident variant calling, ultradeep sequencing coupled with molecular barcoding should be used if sequencing from plasma. WES does not cover all types of genomic aberrations such as fusions or epigenetic mechanisms and this should be borne in mind if initial results are fruitless, and consideration given to including alternative assays such as WGS, RNA-seq or bisulphite sequencing.

6.4 Future directions

The use of ctDNA to guide clinical decisions and drug development is almost certain to increase as more and more supportive data comes to light. There are currently 355 trials worldwide listed as active or recruiting on the US national library of medicine clinical trial database, which gives some idea of the scientific drive. Nevertheless, there is still some distance to go before ctDNA testing becomes everyday reality in the clinic. One major drawback with ctDNA remains the inconsistency in its detection. Methods for ctDNA detection have improved substantially in recent years and the advent of digital PCR has been a real boon. Nevertheless, 24% of patients with active, advanced disease did not have trackable ctDNA in my data. This lack of detection is a common theme across all cancer types and remains a poorly understood phenomenon. It may be due to lack of assay sensitivity, or

other cause, such as variability in tumour shedding of ctDNA. Further work needs to be done to better understand this issue and improve the universality in detection of this liquid biopsy.

Looking to the future, there are several areas in thyroid cancer where current management is contentious and ctDNA could have a positive impact. For example, it could play a role in risk stratification of thyroid nodules, and assist in the longstanding issue with indeterminate nodules, helping to define which patients are most likely to have malignant disease and require surgery. It might also prove useful for decision making with regards the requirement for adjuvant radio-iodine ablation, where there is already a move to de-intensify treatment: lack of detection of ctDNA post-surgery when it was previously detectable might indicate that no radioiodine ablation is needed. Alternatively, for those patients with persistent or recurrent advanced disease it could be used to monitor for disease progression, establish response to targeted therapies, detect resistance and aid in selection of second line therapies. As touched on here the use of plasma ctDNA samples can also be used experimentally to drive exploration of resistance conferring mutations. All of the above will require large scale, prospective studies to confirm their clinical utility, as well as feasibility within a cost-limited healthcare system. However, the step towards truly personalised oncological therapy is one step closer.

References

1. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. Nature Reviews Endocrinology. 2011;7:569.
2. Sobrinho-Simoes M, Eloy C, Magalhaes J, Lobo C, Amaro T. Follicular thyroid carcinoma. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2011;24 Suppl 2:S10-8.
3. Mitchell AL, Gandhi A, Scott-Coombes D, Perros P. Management of thyroid cancer: United Kingdom National Multidisciplinary Guidelines. The Journal of Laryngology & Otology. 2016;130(S2):S150-S60.
4. Nikiforova MN, Nikiforov YE. Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. Expert review of molecular diagnostics. 2008;8(1):83-95.
5. Aschebrook-Kilfoy B, Ward MH, Sabra MM, Devesa SS. Thyroid cancer incidence patterns in the United States by histologic type, 1992-2006. Thyroid : official journal of the American Thyroid Association. 2011;21(2):125-34.
6. Prior IA, Lewis PD, Mattos C. A comprehensive survey of Ras mutations in cancer. Cancer research. 2012;72(10):2457-67.
7. Bhaijee F, Nikiforov YE. Molecular analysis of thyroid tumors. Endocrine pathology. 2011;22(3):126-33.
8. Xing M. BRAF mutation in thyroid cancer. Endocrine-related cancer. 2005;12(2):245-62.
9. Li C, Lee KC, Schneider EB, Zeiger MA. BRAF V600E mutation and its association with clinicopathological features of papillary thyroid cancer: a meta-analysis. The Journal of clinical endocrinology and metabolism. 2012;97(12):4559-70.
10. Xing M, Westra WH, Tufano RP, Cohen Y, Rosenbaum E, Rhoden KJ, et al. BRAF mutation predicts a poorer clinical prognosis for papillary thyroid cancer. The Journal of clinical endocrinology and metabolism. 2005;90(12):6373-9.
11. Liu R, Bishop J, Zhu G, Zhang T, Ladenson PW, Xing M. Mortality Risk Stratification by Combining BRAF V600E and TERT Promoter Mutations in Papillary Thyroid Cancer: Genetic Duet of BRAF and TERT Promoter Mutations in Thyroid Cancer Mortality. JAMA oncology. 2017;3(2):202-8.
12. Jhiang SM, Sagartz JE, Tong Q, Parker-Thornburg J, Capen CC, Cho JY, et al. Targeted expression of the ret/PTC1 oncogene induces papillary thyroid carcinomas. Endocrinology. 1996;137(1):375-8.
13. Powell DJ, Jr., Russell J, Nibu K, Li G, Rhee E, Liao M, et al. The RET/PTC3 oncogene: metastatic solid-type papillary carcinomas in murine thyroids. Cancer research. 1998;58(23):5523-8.
14. Bounacer A, Wicker R, Caillou B, Cailleux AF, Sarasin A, Schlumberger M, et al. High prevalence of activating ret proto-oncogene rearrangements, in thyroid tumors from patients who had received external radiation. Oncogene. 1997;15(11):1263-73.
15. Hamatani K, Eguchi H, Ito R, Mukai M, Takahashi K, Taga M, et al. RET/PTC rearrangements preferentially occurred in papillary thyroid cancer among atomic bomb survivors exposed to high radiation dose. Cancer research. 2008;68(17):7176-82.
16. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. Science (New York, NY). 2013;339(6122):957-9.
17. Liu X, Bishop J, Shan Y, Pai S, Liu D, Murugan AK, et al. Highly prevalent TERT promoter mutations in aggressive thyroid cancers. Endocrine-related cancer. 2013;20(4):603-10.
18. Xing M, Liu R, Liu X, Murugan AK, Zhu G, Zeiger MA, et al. BRAF V600E and TERT promoter mutations cooperatively identify the most aggressive papillary thyroid cancer with highest recurrence. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(25):2718-26.

19. Vu-Phan D, Grachtchouk V, Yu J, Colby LA, Wicha MS, Koenig RJ. The thyroid cancer PAX8-PPARG fusion protein activates Wnt/TCF-responsive cells that have a transformed phenotype. Endocrine-related cancer. 2013;20(5):725-39.
20. Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. Cell. 2014;159(3):676-90.
21. Fagin JA, Wells SA, Jr. Biologic and Clinical Perspectives on Thyroid Cancer. The New England journal of medicine. 2016;375(11):1054-67.
22. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. JAMA oncology. 2016;2(8):1023-9.
23. Chakravarty D, Santos E, Ryder M, Knauf JA, Liao XH, West BL, et al. Small-molecule MAPK inhibitors restore radioiodine incorporation in mouse thyroid cancers with conditional BRAF activation. The Journal of clinical investigation. 2011;121(12):4700-11.
24. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW, 2nd, Tallini G, et al. RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. The Journal of clinical endocrinology and metabolism. 2003;88(5):2318-26.
25. Motoi N, Sakamoto A, Yamochi T, Horiuchi H, Motoi T, Machinami R. Role of ras mutation in the progression of thyroid carcinoma of follicular epithelial origin. Pathology, research and practice. 2000;196(1):1-7.
26. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, et al. PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. Science (New York, NY). 2000;289(5483):1357-60.
27. Elisei R, Cosci B, Romei C, Bottici V, Renzini G, Molinaro E, et al. Prognostic Significance of Somatic RET Oncogene Mutations in Sporadic Medullary Thyroid Cancer: A 10-Year Follow-Up Study. The Journal of Clinical Endocrinology & Metabolism. 2008;93(3):682-7.
28. Ji JH, Oh YL, Hong M, Yun JW, Lee HW, Kim D, et al. Identification of Driving ALK Fusion Genes and Genomic Landscape of Medullary Thyroid Cancer. PLoS genetics. 2015;11(8):e1005467.
29. Frank-Raue K, Rybicki LA, Erlic Z, Schweizer H, Winter A, Milos I, et al. Risk profiles and penetrance estimations in multiple endocrine neoplasia type 2A caused by germline RET mutations located in exon 10. Human mutation. 2011;32(1):51-8.
30. Xu B, Ghossein R. Genomic Landscape of poorly Differentiated and Anaplastic Thyroid Carcinoma. Endocrine pathology. 2016;27(3):205-12.
31. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. The Journal of clinical investigation. 2016;126(3):1052-66.
32. Kunstman JW, Juhlin CC, Goh G, Brown TC, Stenman A, Healy JM, et al. Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. Human molecular genetics. 2015;24(8):2318-29.
33. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
34. CRUK. Incidence Trends Over Time for Common Cancers: Cancer Research UK; 2018 [updated 13 February 2018. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared#heading-Three.
35. Davies L, Welch HG. Current thyroid cancer trends in the United States. JAMA otolaryngology-- head & neck surgery. 2014;140(4):317-22.
36. Grodski S, Brown T, Sidhu S, Gill A, Robinson B, Learoyd D, et al. Increasing incidence of thyroid cancer is due to increased pathologic detection. Surgery. 2008;144(6):1038-43; discussion 43.

37. Jegerlehner S, Bulliard JL, Aujesky D, Rodondi N, Germann S, Konzelmann I, et al. Overdiagnosis and overtreatment of thyroid cancer: A population-based temporal trend study. PloS one. 2017;12(6):e0179387.
38. CRUK. Thyroid Cancer Mortality Trends Over Time: Cancer Research UK; 2018 [updated 13 June 2018. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/thyroid-cancer/mortality#ref.
39. Lubitz CC, Kong CY, McMahon PM, Daniels GH, Chen Y, Economopoulos KP, et al. Annual financial impact of well-differentiated thyroid cancer care in the United States. Cancer. 2014;120(9):1345-52.
40. The Human Genome Project Completion: Frequently Asked Questions: National Human Genome Research Institute; 2010 [updated Oct 2010. Available from: https://www.genome.gov/11006943/human-genome-project-completion-frequently-asked-questions/.
41. Wetterstrand K. DNA Sequencing Costs: Data: National Human Genome Research Institute; 2016 [updated May 2016. Available from: https://www.genome.gov/sequencingcostsdata/.
42. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clinica chimica acta; international journal of clinical chemistry. 2001;313(1-2):139-42.
43. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer research. 2001;61(4):1659-65.
44. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(6):579-86.
45. Reckamp KL, Melnikova VO, Karlovich C, Sequist LV, Camidge DR, Wakelee H, et al. A Highly Sensitive and Quantitative Test Platform for Detection of NSCLC EGFR Mutations in Urine and Plasma. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2016;11(10):1690-700.
46. Li F, Huang J, Ji D, Meng Q, Wang C, Chen S, et al. Utility of urinary circulating tumor DNA for EGFR mutation detection in different stages of non-small cell lung cancer patients. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2017.
47. Mithani SK, Smith IM, Zhou S, Gray A, Koch WM, Maitra A, et al. Mitochondrial resequencing arrays detect tumor-specific mutations in salivary rinses of patients with head and neck cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13(24):7335-40.
48. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008;14(9):985-90.
49. To EW, Chan KC, Leung SF, Chan LY, To KF, Chan AT, et al. Rapid clearance of plasma Epstein-Barr virus DNA after surgical treatment of nasopharyngeal carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research. 2003;9(9):3254-9.
50. Yao W, Mei C, Nan X, Hui L. Evaluation and comparison of in vitro degradation kinetics of DNA in serum, urine and saliva: A qualitative study. Gene. 2016;590(1):142-8.
51. Lo YM, Chan KC, Sun H, Chen EZ, Jiang P, Lun FM, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Science translational medicine. 2010;2(61):61ra91.
52. Kamel AM, Teama S, Fawzy A, El Deftar M. Plasma DNA integrity index as a potential molecular diagnostic marker for breast cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2016;37(6):7565-72.
53. Szpechcinski A, Rudzinski P, Kupis W, Langfort R, Orlowski T, Chorostowska-Wynimko J. Plasma cell-free DNA levels and integrity in patients with chest radiological findings: NSCLC versus benign lung nodules. Cancer letters. 2016;374(2):202-7.

54. Bedin C, Enzo MV, Del Bianco P, Pucciarelli S, Nitti D, Agostini M. Diagnostic and prognostic role of cell-free DNA testing for colorectal cancer patients. International journal of cancer. 2017;140(8):1888-98.
55. Sun K, Jiang P, Chan KC, Wong J, Cheng YK, Liang RH, et al. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(40):E5503-12.
56. Korabecna M, Opatrna S, Wirth J, Rulcova K, Eiselt J, Sefrna F, et al. Cell-free plasma DNA during peritoneal dialysis and hemodialysis and in patients with chronic kidney disease. Annals of the New York Academy of Sciences. 2008;1137:296-301.
57. Destouni A, Vrettou C, Antonatos D, Chouliaras G, Traeger-Synodinos J, Patsilinakos S, et al. Cell-free DNA levels in acute myocardial infarction patients during hospitalization. Acta cardiologica. 2009;64(1):51-7.
58. Beiter T, Fragasso A, Hudemann J, Niess AM, Simon P. Short-term treadmill running as a model for studying cell-free DNA kinetics in vivo. Clinical chemistry. 2011;57(4):633-6.
59. Mandel P, Metais P. Les acides nucléiques du plasma sanguin chez l'homme. Comptes rendus des seances de la Societe de biologie et de ses filiales. 1948;142(3-4):241-3.
60. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737-8.
61. Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. Cancer research. 1977;37(3):646-50.
62. Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M. Neoplastic characteristics of the DNA found in the plasma of cancer patients. Oncology. 1989;46(5):318-22.
63. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, et al. Presence of fetal DNA in maternal plasma and serum. Lancet (London, England). 1997;350(9076):485-7.
64. Lo YM. Fetal DNA in maternal plasma: biology and diagnostic applications. Clinical chemistry. 2000;46(12):1903-6.
65. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 1977;74(12):5463-7.
66. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine. 2012;366(10):883-92.
67. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. The New England journal of medicine. 2017.
68. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature. 2017;545(7655):446-51.
69. Murtaza M, Dawson SJ, Pogrebniak K, Rueda OM, Provenzano E, Grant J, et al. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. Nature communications. 2015;6:8760.
70. Bruggemann M, Kotrova M. Minimal residual disease in adult ALL: technical aspects and implications for correct clinical interpretation. Blood advances. 2017;1(25):2456-66.
71. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Science translational medicine. 2015;7(302):302ra133.
72. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. Science translational medicine. 2016;8(346):346ra92.
73. . !!! INVALID CITATION !!! (60, 61).
74. Marchetti A, Palma JF, Felicioni L, De Pas TM, Chiari R, Del Grammastro M, et al. Early Prediction of Response to Tyrosine Kinase Inhibitors by Quantification of EGFR Mutations in Plasma

of NSCLC Patients. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2015;10(10):1437-43.
75. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. The New England journal of medicine. 2013;368(13):1199-209.
76. Tie J, Kinde I, Wang Y, Wong HL, Roebert J, Christie M, et al. Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. Annals of oncology : official journal of the European Society for Medical Oncology. 2015;26(8):1715-22.
77. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. The New England journal of medicine. 2009;361(10):947-57.
78. Moorcraft SY, Gonzalez de Castro D, Cunningham D, Jones T, Walker BA, Peckitt C, et al. Investigating the feasibility of tumour molecular profiling in gastrointestinal malignancies in routine clinical practice. Annals of oncology : official journal of the European Society for Medical Oncology. 2018;29(1):230-6.
79. Andre F, Bachelot T, Commo F, Campone M, Arnedos M, Dieras V, et al. Comparative genomic hybridisation array and DNA sequencing to direct treatment of metastatic breast cancer: a multicentre, prospective trial (SAFIR01/UNICANCER). The Lancet Oncology. 2014;15(3):267-74.
80. Hauschild A, Grob JJ, Demidov LV, Jouary T, Gutzmer R, Millward M, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. Lancet (London, England). 2012;380(9839):358-65.
81. Planchard D, Smit EF, Groen HJM, Mazieres J, Besse B, Helland A, et al. Dabrafenib plus trametinib in patients with previously untreated BRAF(V600E)-mutant metastatic non-small-cell lung cancer: an open-label, phase 2 trial. The Lancet Oncology. 2017;18(10):1307-16.
82. Baselga J, Im SA, Iwata H, Cortes J, De Laurentiis M, Jiang Z, et al. Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, HER2-negative, advanced breast cancer (BELLE-2): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology. 2017;18(7):904-16.
83. BLU-667 Controls RET-Altered Thyroid Cancers. Cancer discovery. 2019;9(9):Of5.
84. LOXO-292 DEMONSTRATES PROMISING ANTI-TUMOUR ACTIVITY IN RET-ALTERED THYROID CANCER: European Society for Medical Oncology; 2019 [updated 29 Sep 2019. Available from: https://www.esmo.org/oncology-news/loxo-292-demonstrates-promising-anti-tumour-activity-in-ret-altered-thyroid-cancer.
85. Subbiah V, Kreitman RJ, Wainberg ZA, Cho JY, Schellens JHM, Soria JC, et al. Dabrafenib and Trametinib Treatment in Patients With Locally Advanced or Metastatic BRAF V600-Mutant Anaplastic Thyroid Cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2018;36(1):7-13.
86. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nature genetics. 2013;45(12):1439-45.
87. Sefrioui D, Perdrix A, Sarafan-Vasseur N, Dolfus C, Dujon A, Picquenot J-M, et al. Short report: Monitoring ESR1 mutations by circulating tumor DNA in aromatase inhibitor resistant metastatic breast cancer. International journal of cancer. 2015;137(10):2513-9.
88. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, et al. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(21):6494-501.
89. Thompson JC, Yee SS, Troxel AB, Savitch SL, Fan R, Balli D, et al. Detection of therapeutically targetable driver and resistance mutations in lung cancer patients by next generation sequencing of cell-free circulating tumor DNA. Clinical cancer research : an official journal of the American Association for Cancer Research. 2016;22(23):5772-82.

90. Taniguchi K, Uchida J, Nishino K, Kumagai T, Okuyama T, Okami J, et al. Quantitative Detection of EGFR Mutations in Circulating Tumor DNA Derived from Lung Adenocarcinomas. Clinical Cancer Research. 2011;17(24):7808-15.
91. FDA approves first blood test to detect gene mutation associated with non-small cell lung cancer: U.S. Department of Health and Human Services; 2016 [February 2019]. Available from: https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm504488.htm.
92. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Science translational medicine. 2014;6(224):224ra24.
93. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. The New England journal of medicine. 2014;371(26):2488-98.
94. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. The New England journal of medicine. 2014;371(26):2477-87.
95. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126(1):9-16.
96. Lee TH, Montalvo L, Chrebtow V, Busch MP. Quantitation of genomic DNA in plasma and serum samples: higher concentrations of genomic DNA found in serum than in plasma. Transfusion. 2001;41(2):276-82.
97. El Messaoudi S, Rolet F, Mouliere F, Thierry AR. Circulating cell free DNA: Preanalytical considerations. Clinica chimica acta; international journal of clinical chemistry. 2013;424:222-30.
98. Botezatu I, Serdyuk O, Potapova G, Shelepov V, Alechina R, Molyaka Y, et al. Genetic analysis of DNA excreted in urine: a new approach for detecting specific genomic DNA sequences from cells dying in an organism. Clinical chemistry. 2000;46(8 Pt 1):1078-84.
99. Tian F, Liao Y, Zhang Y. Variations in Transrenal DNA and Comparison with Plasma DNA as a Diagnostic Marker for Colorectal Cancer. The International Journal of Biological Markers. 2017;32(4):434-40.
100. Alama A, Truini A, Coco S, Genova C, Grossi F. Prognostic and predictive relevance of circulating tumor cells in patients with non-small-cell lung cancer. Drug discovery today. 2014;19(10):1671-6.
101. Warkiani ME, Khoo BL, Wu L, Tay AKP, Bhagat AAS, Han J, et al. Ultra-fast, label-free isolation of circulating tumor cells from blood using spiral microfluidics. Nature protocols. 2015;11:134.
102. Krebs MG, Hou JM, Ward TH, Blackhall FH, Dive C. Circulating tumour cells: their utility in cancer management and predicting outcomes. Ther Adv Med Oncol. 2010;2(6):351-65.
103. Medina Diaz I, Nocon A, Mehnert DH, Fredebohm J, Diehl F, Holtrup F. Performance of Streck cfDNA Blood Collection Tubes for Liquid Biopsy Testing. PloS one. 2016;11(11):e0166354.
104. Kang Q, Henry NL, Paoletti C, Jiang H, Vats P, Chinnaiyan AM, et al. Comparative analysis of circulating tumor DNA stability In K3EDTA, Streck, and CellSave blood collection tubes. Clinical biochemistry. 2016;49(18):1354-60.
105. Sherwood JL, Corcoran C, Brown H, Sharpe AD, Musilova M, Kohlmann A. Optimised Pre-Analytical Methods Improve KRAS Mutation Detection in Circulating Tumour DNA (ctDNA) from Patients with Non-Small Cell Lung Cancer (NSCLC). PloS one. 2016;11(2):e0150197-e.
106. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European journal of cancer (Oxford, England : 1990). 2009;45(2):228-47.
107. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature Biotechnology. 2013;31:213.

108. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Research. 2016;44(11):e108.
109. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 2010;38(16):e164.
110. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Current protocols in bioinformatics. 2013;43:11.0.1-33.
111. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS computational biology. 2016;12(4):e1004873.
112. Pacini F, Mariotti S, Formica N, Elisei R, Anelli S, Capotorti E, et al. Thyroid autoantibodies in thyroid cancer: incidence and relationship with tumour outcome. Acta endocrinologica. 1988;119(3):373-80.
113. Spencer CA, Takeuchi M, Kazarosyan M, Wang CC, Guttler RB, Singer PA, et al. Serum thyroglobulin autoantibodies: prevalence, influence on serum thyroglobulin measurement, and prognostic significance in patients with differentiated thyroid carcinoma. The Journal of clinical endocrinology and metabolism. 1998;83(4):1121-7.
114. Lee RJ, Gremel G, Marshall A, Myers KA, Fisher N, Dunn J, et al. Circulating tumor DNA predicts survival in patients with resected high risk stage II/III melanoma. Annals of oncology : official journal of the European Society for Medical Oncology. 2017.
115. Reinert T, Scholer LV, Thomsen R, Tobiasen H, Vang S, Nordentoft I, et al. Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery. Gut. 2016;65(4):625-34.
116. Chuang TC, Chuang AY, Poeta L, Koch WM, Califano JA, Tufano RP. Detectable BRAF mutation in serum DNA samples from patients with papillary thyroid carcinomas. Head & neck. 2010;32(2):229-34.
117. Cradic KW, Milosevic D, Rosenberg AM, Erickson LA, McIver B, Grebe SK. Mutant BRAF(T1799A) can be detected in the blood of papillary thyroid carcinoma patients and correlates with disease status. The Journal of clinical endocrinology and metabolism. 2009;94(12):5001-9.
118. Pupilli C, Pinzani P, Salvianti F, Fibbi B, Rossi M, Petrone L, et al. Circulating BRAFV600E in the diagnosis and follow-up of differentiated papillary thyroid carcinoma. The Journal of clinical endocrinology and metabolism. 2013;98(8):3359-65.
119. Lubitz CC, Parangi S, Holm TM, Bernasconi MJ, Schalck AP, Suh H, et al. Detection of Circulating BRAF(V600E) in Patients with Papillary Thyroid Carcinoma. The Journal of molecular diagnostics : JMD. 2016;18(1):100-8.
120. Lubitz CC, Zhan T, Gunda V, Amin S, Gigliotti BJ, Fingeret AL, et al. Circulating BRAFV600E levels correlate with treatment in patients with thyroid carcinoma. Thyroid : official journal of the American Thyroid Association. 2018.
121. Patel KB. Detection of Circulating Thyroid Tumor DNA in Patients with Thyroid Nodules. Electronic Thesis and Dissertation Repository: The University of Western Ontario; 2015.
122. Almubarak H, Qassem E, Alghofaili L, Alzahrani AS, Karakas B. Non-invasive Molecular Detection of Minimal Residual Disease in Papillary Thyroid Cancer Patients. Front Oncol. 2019;9:1510.
123. Kwak JY, Jeong JJ, Kang SW, Park S, Choi JR, Park SJ, et al. Study of peripheral BRAF(V600E) mutation as a possible novel marker for papillary thyroid carcinomas. Head & neck. 2013;35(11):1630-3.
124. Zane M, Agostini M, Enzo MV, Casal Ide E, Del Bianco P, Torresan F, et al. Circulating cell-free DNA, SLC5A8 and SLC26A4 hypermethylation, BRAF(V600E): A non-invasive tool panel for early detection of thyroid cancer. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2013;67(8):723-30.

125. Condello V, Macerola E, Ugolini C, De Napoli L, Romei C, Materazzi G, et al. Analysis of circulating tumor DNA does not improve the clinical management of patients with locally advanced and metastatic papillary thyroid carcinoma. Head & neck. 2018.
126. Cote GJ, Evers C, Hu MI, Grubbs EG, Williams MD, Hai T, et al. Prognostic Significance of Circulating RET M918T Mutated Tumor DNA in Patients With Advanced Medullary Thyroid Carcinoma. The Journal of clinical endocrinology and metabolism. 2017;102(9):3591-9.
127. Sandulache VC, Williams MD, Lai SY, Lu C, William WN, Busaidy NL, et al. Real-Time Genomic Characterization Utilizing Circulating Cell-Free DNA in Patients with Anaplastic Thyroid Carcinoma. Thyroid : official journal of the American Thyroid Association. 2017;27(1):81-7.
128. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Research. 2017;45(D1):D777-D83.
129. Hocevar M, Auersperg M, Stanovnik L. The dynamics of serum thyroglobulin elimination from the body after thyroid surgery. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology. 1997;23(3):208-10.
130. Giovanella L, Ceriani L, Maffioli M. Postsurgery serum thyroglobulin disappearance kinetic in patients with differentiated thyroid carcinoma. Head & neck. 2010;32(5):568-71.
131. Rivas AM, Nassar A, Zhang J, Casler JD, Chindris AM, Smallridge R, et al. ThyroSeq((R))V2.0 Molecular Testing: A Cost-Effective Approach for the Evaluation of Indeterminate Thyroid Nodules. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists. 2018;24(9):780-8.
132. Liu R, Xing M. TERT promoter mutations in thyroid cancer. Endocrine-related cancer. 2016;23(3):R143-55.
133. Mon SY, Riedlinger G, Abbott CE, Seethala R, Ohori NP, Nikiforova MN, et al. Cancer risk and clinicopathological characteristics of thyroid nodules harboring thyroid-stimulating hormone receptor gene mutations. Diagnostic cytopathology. 2018;46(5):369-77.
134. Nicolson NG, Murtha TD, Dong W, Paulsson JO, Choi J, Barbieri AL, et al. Comprehensive Genetic Analysis of Follicular Thyroid Carcinoma Predicts Prognosis Independent of Histology. The Journal of clinical endocrinology and metabolism. 2018;103(7):2640-50.
135. Vandetanib for treating medullary thyroid cancer: NICE; 2018 [cited 2019. Technology appraisal guidance [TA550]]. Available from: https://www.nice.org.uk/guidance/TA550/chapter/1-Recommendations.
136. Asić K. Dominant mechanisms of primary resistance differ from dominant mechanisms of secondary resistance to targeted therapies. Critical Reviews in Oncology/Hematology. 2016;97:178-96.
137. Schlumberger M, Tahara M, Wirth LJ, Robinson B, Brose MS, Elisei R, et al. Lenvatinib versus Placebo in Radioiodine-Refractory Thyroid Cancer. New England Journal of Medicine. 2015;372(7):621-30.
138. Sosman JA, Kim KB, Schuchter L, Gonzalez R, Pavlick AC, Weber JS, et al. Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. The New England journal of medicine. 2012;366(8):707-14.
139. Brose MS, Cabanillas ME, Cohen EEW, Wirth LJ, Riehl T, Yue H, et al. Vemurafenib in patients with BRAFV600E-positive metastatic or unresectable papillary thyroid cancer refractory to radioactive iodine: a non-randomised, multicentre, open-label, phase 2 trial. The Lancet Oncology. 2016;17(9):1272-82.
140. Duquette M, Sadow PM, Husain A, Sims JN, Antonello ZA, Fischer AH, et al. Metastasis-associated MCL1 and P16 copy number alterations dictate resistance to vemurafenib in a BRAFV600E patient-derived papillary thyroid carcinoma preclinical model. Oncotarget. 2015;6(40):42445-67.

141. Montero-Conde C, Ruiz-Llorente S, Dominguez JM, Knauf JA, Viale A, Sherman EJ, et al. Relief of feedback inhibition of HER3 transcription by RAF and MEK inhibitors attenuates their antitumor effects in BRAF-mutant thyroid carcinomas. Cancer discovery. 2013;3(5):520-33.
142. Cheng W, Zhu G, Liu R, Xing M, Wang H. Robust Thyroid Gene Expression and Radioiodine Uptake Induced by Simultaneous Suppression of BRAF V600E and Histone Deacetylase in Thyroid Cancer Cells. The Journal of Clinical Endocrinology & Metabolism. 2016;101(3):962-71.
143. Carlomagno F, Guida T, Anaganti S, Vecchio G, Fusco A, Ryan AJ, et al. Disease associated mutations at valine 804 in the RET receptor tyrosine kinase confer resistance to selective kinase inhibitors. Oncogene. 2004;23(36):6056-63.
144. Isham CR, Netzel BC, Bossou AR, Milosevic D, Cradic KW, Grebe SK, et al. Development and Characterization of a Differentiated Thyroid Cancer Cell Line Resistant to VEGFR-Targeted Kinase Inhibitors. The Journal of Clinical Endocrinology & Metabolism. 2014;99(6):E936-E43.
145. Bazan V, Migliavacca M, Zanna I, Tubiolo C, Grassi N, Latteri MA, et al. Specific codon 13 K-ras mutations are predictive of clinical outcome in colorectal cancer patients, whereas codon 12 K-ras mutations are associated with mucinous histotype. Annals of Oncology. 2002;13(9):1438-46.
146. Branford S, Rudzki Z, Walsh S, Parkinson I, Grigg A, Szer J, et al. Detection of BCR-ABL mutations in patients with CML treated with imatinib is virtually always accompanied by clinical resistance, and mutations in the ATP phosphate-binding loop (P-loop) are associated with a poor prognosis. Blood. 2003;102(1):276-83.
147. Treeck O, Buechler C, Ortmann O. Chemerin and Cancer. International journal of molecular sciences. 2019;20(15):3750.
148. Parolini S, Santoro A, Marcenaro E, Luini W, Massardi L, Facchetti F, et al. The role of chemerin in the colocalization of NK and dendritic cell subsets into inflamed tissues. Blood. 2007;109(9):3625-32.
149. Li JJ, Yin HK, Guan DX, Zhao JS, Feng YX, Deng YZ, et al. Chemerin suppresses hepatocellular carcinoma metastasis through CMKLR1-PTEN-Akt axis. British journal of cancer. 2018;118(10):1337-48.
150. Liu-Chittenden Y, Jain M, Gaskins K, Wang S, Merino MJ, Kotian S, et al. RARRES2 functions as a tumor suppressor by promoting β-catenin phosphorylation/degradation and inhibiting p38 phosphorylation in adrenocortical carcinoma. Oncogene. 2017;36(25):3541-52.
151. Wang C, Wu WK, Liu X, To KF, Chen GG, Yu J, et al. Increased serum chemerin level promotes cellular invasiveness in gastric cancer: a clinical and experimental study. Peptides. 2014;51:131-8.
152. Kaur J, Adya R, Tan BK, Chen J, Randeva HS. Identification of chemerin receptor (ChemR23) in human endothelial cells: chemerin-induced endothelial angiogenesis. Biochemical and biophysical research communications. 2010;391(4):1762-8.
153. Zhao S, Li C, Ye Y-b, Peng F, Chen Q. Expression of Chemerin Correlates With a Favorable Prognosis in Patients With Non-Small Cell Lung Cancer. Laboratory Medicine. 2011;42(9):553-7.
154. Zhang J, Jin HC, Zhu AK, Ying RC, Wei W, Zhang FJ. Prognostic significance of plasma chemerin levels in patients with gastric cancer. Peptides. 2014;61:7-11.
155. Yong Gao MS, Yan Zhao, Hongyang Jiang, Yan Zhang. Serum Chemerin and thyrotropin expression levels in patients with thyroid cancer and roles in predicting survival. Int J Clin Exp Med. 2019;12(6):7561-8.
156. Kataoka K, Ogawa S. Variegated RHOA mutations in human cancers. Experimental hematology. 2016;44(12):1123-9.
157. Cotton S, Azevedo R, Gaiteiro C, Ferreira D, Lima L, Peixoto A, et al. Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. Molecular oncology. 2017;11(8):895-912.
158. Das S, Rachagani S, Torres-Gonzalez MP, Lakshmanan I, Majhi PD, Smith LM, et al. Carboxyl-terminal domain of MUC16 imparts tumorigenic and metastatic functions through nuclear translocation of JAK2 to pancreatic cancer cells. Oncotarget. 2015;6(8):5772-87.

159. Lakshmanan I, Ponnusamy MP, Das S, Chakraborty S, Haridas D, Mukhopadhyay P, et al. MUC16 induced rapid G2/M transition via interactions with JAK2 for increased proliferation and anti-apoptosis in breast cancer cells. Oncogene. 2012;31(7):805-17.
160. Theriault C, Pinard M, Comamala M, Migneault M, Beaudin J, Matte I, et al. MUC16 (CA125) regulates epithelial ovarian cancer cell growth, tumorigenesis and metastasis. Gynecologic oncology. 2011;121(3):434-43.
161. Chen SH, Hung WC, Wang P, Paul C, Konstantopoulos K. Mesothelin binding to CA125/MUC16 promotes pancreatic cancer cell motility and invasion via MMP-7 activation. Scientific reports. 2013;3:1870.
162. Patankar MS, Jing Y, Morrison JC, Belisle JA, Lattanzio FA, Deng Y, et al. Potent suppression of natural killer cell response mediated by the ovarian tumor marker CA125. Gynecologic oncology. 2005;99(3):704-13.
163. Comamala M, Pinard M, Theriault C, Matte I, Albert A, Boivin M, et al. Downregulation of cell surface CA125/MUC16 induces epithelial-to-mesenchymal transition and restores EGFR signalling in NIH:OVCAR3 ovarian carcinoma cells. British journal of cancer. 2011;104(6):989-99.
164. Boivin M, Lane D, Piche A, Rancourt C. CA125 (MUC16) tumor antigen selectively modulates the sensitivity of ovarian cancer cells to genotoxic drug-induced apoptosis. Gynecologic oncology. 2009;115(3):407-13.
165. Karg A, Dinc ZA, Basok O, Ucvet A. MUC4 expression and its relation to ErbB2 expression, apoptosis, proliferation, differentiation, and tumor stage in non-small cell lung cancer (NSCLC). Pathology, research and practice. 2006;202(8):577-83.
166. Carraway KL, Perez A, Idris N, Jepson S, Arango M, Komatsu M, et al. Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. Progress in nucleic acid research and molecular biology. 2002;71:149-85.
167. Nagy P, Friedlander E, Tanner M, Kapanen AI, Carraway KL, Isola J, et al. Decreased accessibility and lack of activation of ErbB2 in JIMT-1, a herceptin-resistant, MUC4-expressing breast cancer cell line. Cancer research. 2005;65(2):473-82.
168. Skrypek N, Duchene B, Hebbar M, Leteurtre E, van Seuningen I, Jonckheere N. The MUC4 mucin mediates gemcitabine resistance of human pancreatic cancer cells via the Concentrative Nucleoside Transporter family. Oncogene. 2013;32(13):1714-23.
169. Qiu L, Wu J, Pan C, Tan X, Lin J, Liu R, et al. Downregulation of CDC27 inhibits the proliferation of colorectal cancer cells via the accumulation of p21Cip1/Waf1. Cell Death & Disease. 2016;7(1):e2074-e.
170. Thu KL, Silvester J, Elliott MJ, Ba-alawi W, Duncan MH, Elia AC, et al. Disruption of the anaphase-promoting complex confers resistance to TTK inhibitors in triple-negative breast cancer. Proceedings of the National Academy of Sciences. 2018;115(7):E1570.
171. Shihab HA, Rogers MF, Gough J, Mort M, Cooper DN, Day INM, et al. An integrative approach to predicting the functional effects of non-coding and coding sequence variation. Bioinformatics. 2015;31(10):1536-43.
172. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(8):2240-7.
173. Zhan T, Rindtorff N, Betge J, Ebert MP, Boutros M. CRISPR/Cas9 for cancer research and therapy. Seminars in Cancer Biology. 2019;55:106-19.
174. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533(7603):420-4.
175. Ma Y, Zhang J, Yin W, Zhang Z, Song Y, Chang X. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nature methods. 2016;13(12):1029-35.
176. Hess GT, Frésard L, Han K, Lee CH, Li A, Cimprich KA, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nature methods. 2016;13(12):1036-42.

177. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. Cell. 2010;141(1):69-80.

Abstract of Associated Publication

European Journal of Cancer 103 (2018) 165–175

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Circulating tumour DNA is a potential biomarker for disease progression and response to targeted therapy in advanced thyroid cancer

D.M. Allin [a,*], R. Shaikh [b], P. Carter [b], K. Thway [c], M.T.A. Sharabiani [d], D. Gonzales-de-Castro [b], B. O'Leary [g], I. Garcia-Murillas [g], S. Bhide [e], M. Hubank [b], K. Harrington [a,e], D. Kim [f], K. Newbold [a,e]

[a] *Radiotherapy and Imaging Division, The Institute of Cancer Research, London, UK*
[b] *Centre for Molecular Pathology, The Royal Marsden Hospital, London, UK*
[c] *Histopathology Department, The Royal Marsden Hospital, London, UK*
[d] *Statistics Unit, The Royal Marsden Hospital, London, UK*
[e] *Head & Neck/Thyroid Oncology Department, The Royal Marsden Hospital, London, UK*
[f] *ENT/Head & Neck Department, St George's Hospital, London, UK*
[g] *Breast Cancer Now Research Centre, The Institute of Cancer Research, London, UK*

Received 1 June 2018; received in revised form 16 August 2018; accepted 19 August 2018

KEYWORDS
Circulating tumour DNA;
Thyroid cancer;
Personalised medicine

Abstract ***Background:*** Conventional biomarkers in thyroid cancer are not disease specific and fluctuate in advanced disease, making interpretation difficult. Circulating tumour DNA (ctDNA) has been shown to be a useful biomarker in other solid tumours. This is a multimutational study of ctDNA over multiple timepoints, designed to test the hypothesis that ctDNA is a potential biomarker in patients with advanced thyroid cancer.

Methods: Mutational analysis of archival tumour tissue was performed using NGS with a targeted gene panel. Custom TaqMan assays were designed for plasma ctDNA testing using digital droplet polymerase chain reaction. Concentrations of detected ctDNA were correlated with the conventional biomarker concentration and axial imaging status defined by the Response Evaluation Criteria in Solid Tumours criteria.

Results: Tumour tissue from 51 patients was obtained, with the following histologies: 32 differentiated (differentiated thyroid cancer [DTC]), 15 medullary (medullary thyroid cancer [MTC]), three poorly differentiated and one anaplastic. NGS analysis detected variants in 42 (82%) of cases. Plasma was assayed for these patients in 190 samples, and ctDNA was detected in 67% of patients. Earlier detection of disease progression was noted in three patients with MTC. In two cases (PTC and ATC), where conventional biomarkers were not detectable,

* *Corresponding author:* Molecular Oncology, The Institute of Cancer Research, 237 Fulham Road, London, SW3 6JB, UK.
E-mail address: davidallin@nhs.net (D.M. Allin).

https://doi.org/10.1016/j.ejca.2018.08.013

Appendix

Chapter 3 study - *Trial protocol*

Title	Molecular screening for patients with advanced thyroid cancers
Anticipated start and end dates	April 2015-April 2016
Study design	Part A: feasibility of detection of circulating tumour DNA (ctDNA) Part B: collection of tumour for development of cell lines and patient derived xenografts (PDX)
Number of patients	TBC
Study Objectives	Primary: 1. To collect blood and tumour tissue for Part A: molecular profiling, including extraction of DNA for sequencing and Part B: for establishment of cell lines and patient derived xenografts (PDX) models of thyroid cancer Secondary: 1. To correlate BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARY, β-catenin, p53, PTEN and PI3K mutations in ctDNA with Formalin Fixed Paraffin Embedded tumour tissue (FFPE) mutational analysis. 2. To correlate the quantity of ctDNA fragments 3 monthly in patients with advanced disease on routine follow up with conventional tumour markers (Tg, calcitonin, CEA). 3. To correlate the quantity of ctDNA fragments 3 monthly once relapse suspected by conventional methods (Tg, Calcitonin, CEA, radiological) 4. To correlate the quantity of ctDNA fragments with response (RECIST, Tg, Calcitonin and CEA) to targeted therapies 5. To assess prognostic significance of ctDNA levels in metastatic/advanced thyroid cancer 6. To isolate live tumour cells for studies of novel treatment strategies (combinatorial treatments), therapy resistance and thyroid cancer biology 7. Retrieval and analysis of archival primary tissue blocks for comparison with metastatic tumour sites 8. To collect blood for analysis of non-invasive tests of tumour phenotype
Primary endpoint	Proportion of patients with detectable ctDNA (detectable is defined as any trace of ctDNA in the blood sample).
Secondary and exploratory endpoints	Secondary End points:

	1. To investigate the reproducibility of quantification of ctDNA (2 samples at baseline and in case of relapse) in advanced thyroid cancer 2. To investigate association between BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARϒ, β-catenin, p53, PTEN and PI3K mutations in ctDNA and FFPE BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARϒ, β-catenin, p53, PTEN and PI3K mutational analysis. 3. To correlate the quantity of ctDNA fragments 3 monthly in patients on routine follow up with conventional tumour markers (Tg, calcitonin, CEA, RECIST as appropriate). Exploratory end points 1. To correlate ctDNA fragments with conventional markers of response (RECIST, Tg, Calcitonin and CEA) to targeted therapies 2. To establish cell lines and PDX
Inclusion criteria	• T3-4, N0-1b, M0-M1 thyroid cancer (papillary, follicular, poorly differentiated, anaplastic and medullary thyroid carcinoma). • Patient able to provide informed consent • Patient attending the Royal Marsden Hospital for inpatient or outpatient review • Sufficient tissue sample available to perform analysis (archived or fresh) • Adults >16 years old
Exclusion criteria	• Patients unable to provide informed consent • Thyroid lymphoma • Metastases to the thyroid • No histological confirmed diagnosis • Only cytology available

Study design: This is a biological study to obtain tissue and blood samples from patients with advanced thyroid cancer

Study Rationale

This is a two part study; part A proposes to collect plasma samples to examine how circulating tumour DNA markers correlate with detection of recurrent disease, response to therapy, clinical outcome and pathological data. Archived formalin fixed tumour will be retrieved where possible for comparison; part B aims to use tissue obtained from biopsies of recurrent disease to establish cell lines and tumour explants to further investigate the biology of thyroid cancer in the pre-clinical setting.

Study Objectives:

- **Primary**:

  - To collect blood and tumour tissue for Part A: molecular profiling, including extraction of DNA for sequencing, RNA for expression levels and to identify expressed fusion genes, and proteins for proteomic studies and Part B: for establishment of cell lines and patient derived xenograft (PDX) models of thyroid cancer

- **Secondary:**

  - To correlate BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARY, β-catenin, p53, PTEN and PI3K mutations in ctDNA with Formalin Fixed Paraffin Embedded tumour tissue (FFPE) mutational analysis.
  - To correlate the quantity of ctDNA fragments 3 monthly in patients with advanced disease on routine follow up with conventional tumour markers (Tg, calcitonin, CEA).
  - To correlate the quantity of ctDNA fragments 3 monthly once relapse suspected by conventional methods (Tg, Calcitonin, CEA, radiological)
  - To correlate the quantity of ctDNA fragments with response (RECIST, Tg, Calcitonin and CEA) to targeted therapies
  - To assess prognostic significance of ctDNA levels in metastatic/advanced thyroid cancer
  - To isolate live tumour cells for studies of novel treatment strategies (combinatorial treatments), therapy resistance and thyroid cancer biology
  - Retrieval and analysis of archival primary tissue blocks for comparison with metastatic tumour sites
  - To collect blood for analysis of non-invasive tests of tumour phenotype

Study End points:

- **Primary End point**:
  Proportion of patients with detectable ctDNA (detectable is defined as any trace of ctDNA in the blood sample).

- **Secondary End point:**
  1. To investigate the reproducibility of quantification of ctDNA (2 samples at baseline and in case of relapse) in advanced thyroid cancer

  2. To investigate association between BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARY, β-catenin, p53, PTEN and PI3K mutations in ctDNA and FFPE BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARY, β-catenin, p53, PTEN and PI3K mutational analysis.

  3. To correlate the quantity of ctDNA fragments 3 monthly in patients on routine follow up with conventional tumour markers (Tg, calcitonin, CEA, RECIST as appropriate).

- **Exploratory end point;**
  1. To correlate ctDNA fragments with conventional markers of response (RECIST, Tg, Calcitonin and CEA) to targeted therapies
  2. To establish cell lines and PDX

Study Population:

This is a biological research study involving the collection of blood and tumour tissue.

Inclusion:

- T3-4, N0-1b, M0-M1 thyroid cancer (papillary, follicular, poorly differentiated, anaplastic and medullary thyroid carcinoma).
- Patient able to provide informed consent
- Patient attending the Royal Marsden Hospital for inpatient or outpatient review
- Sufficient tissue sample available to perform analysis (archived or fresh)
- Adults >16 years old

Exclusion:

- Patients unable to provide informed consent
- Thyroid lymphoma
- Metastases to the thyroid
- No histological confirmed diagnosis
- Only cytology available

Output

Part A: A successful feasibility study will allow the development of a multicentre study

Part B: the successful establishment of thyroid cancer cell lines and patient derived xenografts for investigation of targeted therapies and mechanisms of therapy resistance.

Patient identification and screening

Routine clinicopathological data including treatment details will be collected in linked-anonymised fashion for correlation, and held in a secure database.
Repeat research serum samples will be taken at 3 monthly intervals alongside routine serum samples carried out according to clinical protocol.

Eligible patients will be identified at the thyroid unit MDT. Patients will be approached for consideration of study entry whilst attending the oncology outpatient clinic at the Royal Marsden Hospital. Patients will receive a copy of a Patient Information Sheet (PIS) for the study. Patients that meet the eligibility criteria will be recruited and given a unique numerical

identifier. Medications, dose modifications of systemic therapy and supportive treatment will be as per routine clinical practice.

We aim to recruit all eligible patients treated within the Royal Marsden Hospital Thyroid Unit over the next 12 months. We will aim to recruit 53 patients.

Consent

Patients will receive a PIS and will have the opportunity to discuss the study in detail prior to giving consent. Patients may consent on the same working day as receipt of the PIS. Written informed consent will be obtained prior to study entry. Consent procedures will conform to GCP, national and local regulations.

A unique identifier will be allocated to each participant by the trial co-ordinator. Samples will be labelled with patient's unique identifier, timing of sample and date sample obtained.

Following informed consent, patients will have two blood samples obtained at enrolment in the study. This will be taken at the same time as routine blood sample.

Patients will also be asked for consent to a biopsy of newly recurrent or progressing disease if biopsy is feasible or to additional research biopsies being taken at the same time as routine biopsies and to consent to access archival tissue.

The trial will be conducted in compliance with the protocol and Good Clinical Practice.

Subject Withdrawal Criteria

Participation in this study is entirely voluntary and patients can withdraw their consent at any time.

If a subject withdraws from the study prior to obtaining a first sample, the subject will be replaced. Should a patient withdraw following sample collection, any biological or clinical data collected during the study may still be analysed unless the patient expressly requests that this does not occur.

There is no additional follow-up required for withdrawn patients other than routine clinical follow-up.

Study Organisation, Trial Monitoring and Management Strategy

Overall Responsibility

The Chief Investigator will take primary responsibility for the conduct of the study in accordance with the World Medical Association Declaration of Helsinki and subsequent amendments, and the conduct will conform to ICH Good Clinical Practice guidelines and the Research Governance Framework Guidelines.

Data Acquisition, Monitoring, Storage and Analysis

Eligibility criteria will be reviewed by the investigator and the trial co-ordinator. Clinical data with be collected on the Case Report Form (CRF) and entered into the database by the trial team. The database will be monitored periodically to ensure completeness and accuracy.

Management Strategy

The monitoring of the trial will be performed by the study team, study co-ordinator and statistician.

Potential Conflicting Studies

Nil

Study Procedures

Baseline clinical data will be recorded at enrolment. The blood and tissue samples for Part A of the protocol will be forwarded to the RMH Tissue Bank under Janine Salter for collection and preparation then transferred to Centre for Molecular Pathology under David Gonzalez de Castro for analysis. For Part B of the study, tissue samples will be forwarded to the Institute of Cancer Research under Malin Pederson for establishment of cell lines and PDXs (as per the ICR Transport and Receipt of Hazardous Materials guidelines). The potential risk to human subjects is minimal, as samples will be obtained by routine phlebotomy and biopsy practices.

Blood sample and Biopsy collection schedule

If anatomically accessible patients will be asked to consent to a biopsy of the progressing disease at the time of progression and on progression whilst on a targeted therapy.

1. Tumour samples will be obtained at surgery or by radiologically-guided (ultrasound or CT) biopsy. We have detailed standard operating procedures (SOPs), (Appendix 1) in place for this process – including availability of research nurses to transfer the sample promptly in tissue culture medium on ice from the operating theatre or radiology department to RMH tissue bank (Part A) and the research laboratory at ICR (Part B).
2. **Part A:**
   Samples will be processed under the SOP of the Tissue Bank (Appendix 2) and analysed for mutations under the SOP for the CMP (Appendix 3). We will focus on identifying specific genetic abnormalities (so-called "actionable mutations") that may be amenable to modulation by existing or novel targeted therapies. Figure 1.
3. **Part B:**
   i) Tumour samples will be processed according to ICR SOP (Appendix 4). The samples will be disaggregated in tissue culture medium (containing anti-bacterial and anti-fungal antibiotics) and grown *in vitro* for subsequent analyses. The aim of these experiments will be to establish long-term cultures of tumour cells from individual patients and to use them in therapeutic trials *in vitro*. We estimate that we will succeed in establishing tumour cell cultures in approximately 33% of samples (depending on tumour histology).

   ii) Part of the tumour sample will be implanted directly into NOD/SCID or nude mice in an

attempt to establish *in vivo* tumour models derived directly from the patient's tumour. We estimate having sufficient material to attempt tumour culture *in vivo* in up to 15 patients. Once established, such tumours can be used in therapeutic studies (including biological and radiological investigations to define potential biomarkers of activity). We have experience with the technique of establishing tumour xenografts from patient-derived specimens and have recently employed a scientist (with significant expertise in this area) who will be dedicated to optimising this process. The objective of this section of the project will be to generate groups of mice bearing patient-derived tumours and to subject them to randomised trials of specific therapies (using a so-called "mouse hospital" approach). We estimate that we will successfully establish xenograft tumours in 40-60% of cases – and this should allow us an opportunity to study therapeutic responses in multiple different tumours of the same histological type.

Research blood and tissue collection schedule

Blood samples ± tissue from biopsies or archival FFPE tissue will be collected at baseline following study entry and consent.

Repeat blood samples will be taken every 3 months at same time as routine blood samples or on clinical progression. Repeat biopsies will be taken if possible at the time of clinical progression.

Planned specimen analysis

Part A:

Tumour will be identified within the formalin fixed paraffin-embedded (FFPE) tissue samples by a consultant pathologist (KT) and DNA will be isolated using micro dissection. This reference tumour DNA will be analysed (CMP team, D.G-d-C) using a customised or the gastro-intestinal NSG panel. The panel is based on Illumina's TrueSight technology which interrogates both strands of the DNA simultaneously and therefore making it applicable to DNA extracted from FFPE tissue, reducing significantly the risk of false positive results due to deamination artefacts. Samples will be sequenced to an average depth of >1,000x per base pair, allowing the detection of mutations present in at least 5% of alleles. Analysis will be performed using Variant Studio v2.

A digital PCR assay will then be developed based on the findings of the NGS panel to detect the specific mutation or mutations identified in the individual patient's tumour in the plasma. Custom TaqMan primer-probes will be used on the QX100 droplet digital PCR machine. DNA will be extracted from the plasma using a standard, commercially available Qiagen Circulating Nucleic Acid kit within the Tumour bank team (JS).

Plasma and tumour DNA will be analysed using deep sequencing for BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARϒ, β-catenin, p53, PTEN and PI3K mutations using digital PCR.

1ml of the initial blood sample will be used to obtain germ-line sequence information. Th germ-line sequence information will be used only to establish which variants are restricted to the tumour (i.e. are somatic mutations). Germline genetic variants will not be evaluated and will not be reported back to the patient. In the unlikely event that a previously described mutation found in a tumour is found to be a germ-line mutation this information will be fed back to the study CI and the treating clinician.

Part B:

Following surgical excision, tumour samples in excess of that required for diagnostic purposes, will immediately be separated from surrounding tissue by the surgeon (DK) and manually dissected into tissue fragments.

Collection of fresh tumour tissue

1. Three-dimensional organoid models and cell lines

Cell lines from patient samples will be established from fresh thyroid tumor biopsies using established enzymatic and mechanical digestion protocols. In short, tumour fragments will be put in sterile transport tubes filled with 20mL of transport medium enriched with antibiotics. All tissue processing following initial dissection will be performed in a biosafety cabinet (hood) using Biosafety Level 2 (BSL2) techniques in the laboratory. Necrotic parts will be removed if present. Tumour fragments will be processed by chopping with a scalpel or a razor into 50-200µm small pieces, followed by enzymatic digestion (using 1.25mg/mL collagenase type A, 10mg/mL Dispase II or other proteases, in a 37C incubation shaker). These tumour fragments will create tumour organoids that can be used for 3D culturing *in vitro.* Tumour organoids can subsequently be digested with trypsin/EDTA to generate a single cell suspension for 2D cell culturing or FACS *in vitro*. We estimate that the success rate to derive cell lines that grow *in vitro* will be ~30%.

2. Patient derived xenograft models (PDX)

Two pieces of the tumour will go to establishment of PDXs. Necrotic parts of the tumour will be removed and 5x5x5 mm pieces will be implanted subcutaneously in the right flank of 5-6 weeks old female Balb-c nude or Il-2 NGS mice. When the PDXs reach 1500 mm^3 volume, they will be excised, and viable tissue dissected into 5 mm cubes and transplanted into additional mice using the same procedure. Genomic and histological analyses can be used to confirm that the tumors at each point are derived from the starting material.

Collection of frozen tumour tissue

One piece of the sample will be placed in a cryovial and snap-frozen in liquid nitrogen. Frozen tumour tissue will be stored in -80C freezers in the laboratory of Professor Kevin Harrington at the Institute of Cancer Research. Samples will be micro-dissected at the Breakthrough histopathology laboratory and DNA and RNA will be extracted using commercial kits (Qiagen). 1ug of DNA and 1ug of RNA will be used for next generation sequencing techniques (WES; whole exome sequencing and RNA Sequencing), performed within the TPU (Tumour Profiling Unit) at the ICR. DNA and RNA extracted from blood will be used as matched normal control. If there is enough tissue, another piece of the sample will be frozen in a mixture of 90% Fetal Bovine Serum (FBS) / 10% DMSO in a cryovial and stored in a -80C freezer. This sample can later be used to re-implant in mice in case the implanted mice get sick and die or for establishment of cell lines in *in vitro* cultures in case the first attempt is unsuccessful.

Specimen labelling and storage

Part A: samples retained as part of the study will be coded and stored securely within the RMH Tissue bank under the patient unique identifier. Other patient identifiers will be removed prior to storage to maintain confidentiality. The principal investigator and co-investigators in the study will have access to the patient baseline data and the clinical and outcome data to enable correlation of laboratory results with this data. Electronic data analyses will be password protected.

Part B: Samples retained as part of the study will be coded and stored securely in freezers at the Institute of Cancer Research under the patient unique identifier. Other patient identifiers will be removed prior to storage to maintain confidentiality. The principal investigator and co-investigators in the study will have access to the patient baseline data and the clinical and outcome data to enable correlation of laboratory results with this data. Electronic data analyses will be password protected.

Clinical and demographic data collection

The study CRF will be used to collect a summary of the key clinical data regarding each patient enrolled in the study. This will be forwarded to the study clinical fellow or research nurse and the data will be entered into a secure CRS database at the Royal Marsden using the patient's unique identifier to enable correlation with the clinical data whilst maintaining patient confidentiality.

Clinical data to be collected:

Please refer to the Case Report form (CRF)

Study completion and end of trial definition

The end of the pilot study accrual will be when 53 patients have been recruited and the last blood/tissue sample taken as per trial protocol.

Statistical Considerations

Sample Size

This is a feasibility study and if the success rate (*i.e*. ctDNA is detected) is <50% the approach would be considered not feasible however if the success rate is >70% a bigger study will be planned to investigate the prognostic value of ctDNA in thyroid cancer. Therefore, the number required for the feasibility phase would be 53 patients. Using an alpha value of 5% and power of 90% gives a trial size of 53 with a cut-off of 33 patients having detectable ctDNA.to accept the hypothesis.(Ahern et al, Statist. Med 2001; 20:859-866)

Statistical analysis

All quantitative data will be presented as number of observations, mean, standard deviation, minimum and maximum values. If the data cannot be assumed to be normally distributed, following visual inspection, median will be used together with 25% and 75% quartiles. Qualitative data will be presented as number of observations and percentages. When appropriate, data will be presented together with 95% confidence intervals. All statistical tests will be two-tailed and a p-value <0.05 (5%) will be considered as statistically significant.

Primary end point:

The number of patients with detectable ctDNA will be reported as proportions and 95% confidence intervals will be presented, where detectable is defined as any trace of ctDNA in the blood sample.

Secondary end points:

- The reproducibility of quantification of ctDNA from 2 samples at baseline and at relapse will be analysed using a Bland-Altman plot.
- The association between BRAF V600E, RAS, RET/PTC, RET, PAX8/PPARϒ, β-catenin, p53, PTEN and PI3K mutations in ctDNA and FFPE will be analysed using a chi-square test or Fisher exact test as appropriate. Also the concordance of these mutations in ctDNA and FFPE will be presented.
- The association of the quantity of ctDNA fragments (using all timepoints) with conventional tumour markers Tg, calcitonin and CEA will be presented using correlation coefficient. The levels of ctDNA will be plotted against the Tg, calcitonin and CEA.
- The mean change from baseline in ctDNA levels at each 3 month assessment will be plotted by those who did and did not have a Progressive Disease recorded (by RECIST criteria) in the next 3 months. Patients who progressed will be excluded in the following 3 month assessments.
- The association of ctDNA fragments with conventional tumour markers of response (RECIST, Tg, Calcitonin, and CEA) to the different targeted therapies will be summarized using descriptive statistics.

Regulatory & Ethics Committee Approval

Ethical Considerations

The study protocol and any subsequent amendments will be submitted to the Research Ethics Committee (REC). The study will only open to recruitment following approval from both the REC and the Royal Marsden Research and Development department and will conducted as per the ratified protocol. All patient and volunteer information sheets and consent forms will be approved by the REC.

Patients will receive written and verbal information about the study prior to enrolment and will be able to enrol on the same working data as information is received. Risks and potential benefits of study participation will be fully explained. Informed consent will be recorded on an ethically approved consent form. Trained personnel including research nurses can take consent.

Participation in this study is entirely voluntary and patients can withdraw their consent at any time. If a subject withdraws from the study prior to obtaining a first blood test, the subject will be replaced. Should a patient withdraw following sample collection, any biological or clinical data collected during the study may still be analysed and retained in an anonymised form unless the patient expressly requests that this does not occur.

Samples retained as part of the study will be stored securely in freezers at the Institute of Cancer Research under the patient unique identifier. Other patient identifiers will be removed prior to storage to maintain confidentiality. The trial co-ordinator will have access to the patient baseline data

and the clinical and outcome data to enable correlation of laboratory results with this data and maintain confidentiality. Electronic data analyses will be password protected.

There will be no financial incentive to enrol in any part of the study.

The Chief Investigator will update the ethics committee of any new information relating to the study where appropriate.

Data Handling and Record Keeping

Data will be collected and maintained in accordance to ICH-GCP standards. The study CRF will be used to collect a summary of the key clinical data regarding each patient enrolled in the study. This data will be entered into a password-protected database using the patient's unique identifier to enable correlation with the clinical data whilst maintaining patient confidentiality. The CRF will be updated using the electronic patient record (EPR). Any subsequent data analyses and laboratory results will be stored securely on a password-protected drive on the RM or ICR network. The CRFs will be stored in a secure room in the Royal Marsden.

The Chief Investigator will ensure that staff complies with the study protocol. Original consent forms will be signed by the patient and trained staff and will be retained in a secure location in the Royal Marsden. Sufficient data will be recorded to allow linkage of the electronic patient record, study data and laboratory data. All protocol related documentation will be available for GCP audit and inspection as required.

Financing, Indemnity & Insurance

Indemnity is as per usual NHS indemnity arrangements.

Funding for clinical personnel costs and sample procurement for this study will be applied for from the Royal Marsden Hospital NIHR BRC funding and Oracle Cancer Charity. No individual payments per patient will be made to any study or laboratory personnel.

Publication Policy

All publications and presentations relating to the study will be authorised by the chief investigator.