# MOLECULAR ANALYSIS OF CIRCULATING CELL-FREE DNA IN LUNG CANCER

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

**Ricky Mark Trigg**

BSc Honours (University of Leicester)

Department of Cancer Studies

University of Leicester

November 2016

# Abstract

**Molecular analysis of circulating cell-free DNA in lung cancer**

Ricky Mark Trigg

Two-thirds of patients with non-small cell lung cancer (NSCLC) are diagnosed with incurable disease, reflecting the need for early detection. However, research in this area has been hampered by the inability to identify individuals with early-stage tumours or preneoplastic lesions that may progress to malignancy.

In this thesis, the $Kras^{+/LSL-G12D}$ mouse was used to model lung preneoplasia, permitting analysis of circulating-free DNA (cfDNA) in comparison with tumour burden. Using end-point polymerase chain reaction (PCR), tumour-derived cfDNA (ctDNA) was detected in 3/7 $Kras^{+/LSL-G12D}$ mice with lung preneoplasia but no evidence of malignant transformation. Moreover, total levels of cfDNA were elevated in $Kras^{+/LSL-G12D}$ mice (mean, 66.6 ng/mL) relative to $Kras^{+/+}$ mice (mean, 2.5 ng/mL) ($P = 0.006$). To permit a quantitative analysis of cfDNA, digital and quantitative real-time PCR strategies were then developed to detect $Kras^{LSL-G12D}$ and $Kras^{Lox-G12D}$. Using these assays, analysis of the sensitivity of ctDNA in $Kras^{+/LSL-G12D}$ mice may inform its diagnostic potential in human NSCLC.

In a separate project, somatic copy number alteration (SCNA) in cfDNA was investigated in 143 patients with advanced NSCLC. Analysis of SCNA in five commonly amplified genes (*EGFR*, *FGFR1*, *HER2*, *MET* and *PIK3CA*) by qPCR and ddPCR identified gene amplification in 18 patients. Next-generation sequencing identified mutations in 28 patients, of which only 2 patients also had evidence of SCNA. Interestingly, patients with SCNA in cfDNA had significantly poorer overall survival (OS) than patients without (HR, 1.78; 95% CI, 1.05 – 3.02; $P = 0.03$). This significance was improved further when SCNA and mutation data were combined; patients with one or both types of alteration had significantly poorer progression-free survival (PFS) (HR, 1.50; 95% CI, 1.01 – 2.21; $P = 0.02$) and OS (HR, 1.88; 95% CI, 1.24 – 2.84; $P = 0.003$) than patients without. This is the first time SCNA in cfDNA has demonstrated predictive potential in any solid cancer type.

# Acknowledgements

I wish to extend my deepest thanks to my supervisors, Professors Jacqui Shaw and Catrin Pritchard, for their continuous enthusiasm, kind support and expert guidance. Their belief in my ability to complete this endeavour is very important to me, and I greatly appreciate the opportunity to work in an exciting and rapidly advancing area of cancer research.

I would like to thank the other members of my Thesis Committee, Drs Don Jones, Howard Pringle and Alessandro Rufini, for their kind words of encouragement, constructive criticism and enlightening discussions. Dr Pringle shared much of his PCR expertise, and our academic discussions always left me inspired and with refreshed enthusiasm.

I am indebted all members of the JS lab for their support with my laboratory work and their encouragement when things didn't go to plan. In particular, I'd like to thank Drs David Guttery and Barbara Ottolini for their academic input, and Mrs Lindsay Primrose for her kindness and generous technical support, especially during the early stages of my work. I am also grateful to Mrs Susan Giblett and staff within the Division of Biomedical Services for their time and hard work in manipulating the mouse models.

Thank you to Professor Andrew Fry, my undergraduate personal tutor and third-year project supervisor, for allowing me to realise my true potential as an undergraduate student and for encouraging me to seek a PhD position.

Finally, thank you to my family and friends for their unconditional love, support and encouragement throughout my degree and during times of difficulty; I hope to make you proud of me.

I wish to acknowledge the Medical Research Council for their generous funding that has enabled me to pursue this PhD.

# Related Publications

1. Factors that influence quality and yield of circulating-free DNA: a systematic review of the methodology literature [in draft]. Trigg, R.M. and Shaw, J.A.

2. Circulating-free DNA profiling of non-small-cell lung cancer using Ion Torrent next-generation sequencing [in draft]. Second author.

3. Correlating the plasma genotype with outcomes in a phase II trial of docetaxel combined with the Hsp90 inhibitor ganetespib (GALAXY-1) [in review at JAMA Oncology]. Named author.

# Contents

# List of tables

# List of figures

# Abbreviations

| Acronym | Definition |
|---------|------------|
| **AAH** | atypical adenomatous hyperplasia |
| **AC** | adenocarcinoma (lung) |
| **AIS** | adenocarcinoma *in situ* (lung) |
| ***ALK*** | anaplastic lymphoma kinase |
| **ATP** | adenosine triphosphate |
| **bp** | base pair |
| **BLAST** | Basic Local Alignment Search Tool |
| ***BRAF (Braf)*** | v-raf murine sarcoma viral oncogene homologue B1 |
| **cfDNA** | circulating free DNA |
| **CI** | confidence interval |
| **CIS** | carcinoma *in situ* (lung) |
| **CNV** | copy number variation |
| **c-SCLC** | combined small cell lung cancer |
| **CT** | computed tomography |
| **$C_T$** | cycle threshold |
| **CTC** | circulating tumour cell |
| **ctDNA** | circulating tumour DNA |
| **DII** | DNA integrity index |
| **DMSO** | dimethyl sulfoxide |
| **DNA** | deoxyribonucleic acid |
| **dNTP** | deoxynucleotide triphosphate |
| **dsDNA** | double stranded DNA |
| **dPCR** | digital PCR |
| **E** | efficiency (of PCR amplification) |
| **ED** | extensive-stage disease |
| **EDTA** | ethylenediaminetetraacetic acid |
| ***EGFR*** | epidermal growth factor receptor |
| ***EML4*** | echinoderm microtubule-associated protein-like 4 |
| **EpCAM** | epithelial cell adhesion molecule |

| | |
|---|---|
| ***ERBB2*** | human epidermal growth factor receptor 2 (*HER2*) |
| **ESC** | embryonic stem cell |
| **FBS** | fetal bovine serum |
| **FDR** | false discovery rate |
| ***FGFR1*** | fibroblast growth factor receptor 1 |
| **g** | gram |
| **gDNA** | genomic DNA |
| **GEM** | genetically engineered mouse |
| **FFPE** | formalin-fixed, paraffin-embedded |
| **HEPES** | 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid |
| **HET** | heterozygous |
| **H/E** | haematoxylin and eosin |
| **hgDNA** | human genomic DNA |
| **HR** | hazard ratio |
| **IAA** | isoamyl alcohol |
| **IMS** | industrial methylated spirits |
| **kb** | kilobase |
| ***KRAS* (*Kras*)** | Kirsten rat sarcoma viral oncogene homolog |
| **LB** | Luria-Broth |
| **LCC** | large cell carcinoma (lung) |
| **LD** | limited-stage disease |
| **LOH** | loss of heterozygosity |
| **LSL** | Lox-STOP-Lox |
| **M** | molar |
| **MEF** | mouse embryonic fibroblast |
| ***MET*** | hepatocyte growth factor receptor |
| **mg** | milligram |
| **MGB/NFQ** | minor groove binder/ non-fluorescent quencher |
| **MIA** | minimally invasive adenocarcinoma (lung) |
| **mL** | millilitre |
| **mM** | millimolar |
| **mol** | mole |

| | |
|---|---|
| *MYC* | c-Myc |
| *MYCL1* | L-Myc |
| *MYCN* | N-Myc |
| **NCI** | National Cancer Institute |
| **ng** | nanogram |
| **NGS** | next-generation sequencing |
| **nM** | nanomolar |
| *NRAS* | neuroblastoma rat sarcoma viral oncogene homolog |
| **NSCLC** | non-small cell lung carcinoma |
| **NTC** | no template control |
| **OS** | overall survival |
| **p** | short chromosome arm |
| **P** | P-value |
| **PA** | polyadenylation |
| **PBS** | phosphate buffered saline |
| **PCR** | polymerase chain reaction |
| **pen/strep** | penicillin/streptomycin |
| **PFS** | progression-free survival |
| **PFU** | plaque-forming unit |
| **pg** | picogram |
| *PIK3CA* | phosphoinositide-3-kinase, catalytic, alpha polypeptide |
| **pM** | picomolar |
| *PTEN* | phosphatase and tensin homolog |
| **q** | long chromosome arm |
| *q* | FDR-corrected significance level |
| **qPCR** | quantitative PCR |
| **$R^2$** | regression coefficient |
| *RB1* | retinoblastoma protein |
| **rcf** | relative centrifugal force |
| *RET* | rearranged during transfection |
| **RNA** | ribonucleic acid |
| *ROS1* | c-ros oncogene 1, receptor tyrosine kinase |

| | |
|---|---|
| **rpm** | revolutions per minute |
| **RPMI** | Roswell Park Memorial Institute |
| *RPPH1* | ribonuclease P RNA component H1 |
| **RQ** | relative quantification (comparative $C_T$ method) |
| **RT** | room temperature |
| **RTK** | receptor tyrosine kinase |
| **SA** | splice acceptor |
| **SCLC** | small cell lung carcinoma |
| **SCNA** | somatic copy number alteration |
| **SD** | standard deviation |
| **SDS** | sodium dodecyl sulfate |
| **SNV** | single nucleotide variant |
| *SOX2* | (sex determining region Y)-box 2 |
| **SqCC** | squamous cell carcinoma (lung) |
| **$T_a$** | annealing temperature |
| **TE** | Tris-EDTA |
| **TSG** | tumour suppressor gene |
| **TKI** | tyrosine kinase inhibitor |
| **$T_m$** | melting temperature |
| **TNM** | tumour-node-metastasis |
| *TP53* | tumour protein p53 |
| **UV** | ultraviolet |
| **μCT** | micro-computed tomography |
| **μg** | microgram |
| **μL** | microlitre |
| **μM** | micromolar |
| **VIC®** | 2'-chloro-7'-phenyl-1,4-dichloro-6-carboxyfluorescein |
| **v/v** | volume/volume |
| **WT** | wild-type |
| **ºC** | degrees Celsius |
| **6-FAM™** | 6-carboxyfluorescein |
| **$2^{-\Delta\Delta C_T}$** | comparative $C_T$ method (relative quantification) |

# Chapter 1

# Introduction

## 1.1        The human lung

The lungs are located within the thoracic cavity on either side of the mediastinum. The left lung comprises two lobes and the right lung comprises three lobes. An indentation between the two lobes of the left lung, called the cardiac notch, accommodates the heart. The principal function of the lungs is rapid respiratory exchange, which is facilitated by a large surface area-to-volume ratio and a rich blood supply. Air is inhaled through the trachea, the primary bronchi, a series of successively narrowing airways (secondary bronchi, tertiary bronchi, bronchioles, terminal bronchioles, respiratory bronchioles), and finally into the functional units of gaseous exchange; the alveoli (Murray, 2010). The bronchi are characterised by thick muscular walls surrounded by cartilage rings, whilst the bronchioles have thinner walls and no cartilage support (Sun *et al.*, 2007). The epithelium of the major bronchi comprises three differentiated cell types: ciliated cells, goblet (mucous-producing) cells and pulmonary neuroendocrine cells (Fig 1.1). The putative stem cell for these cell types is the basal cell (Boers *et al.*, 1998). In contrast, the epithelia of the respiratory bronchioles and alveoli are characterised by Clara cells and type II pneumocytes, respectively, both of which are thought to arise from a common precursor; the bronchioalveolar stem cell (Kim *et al.*, 2005).
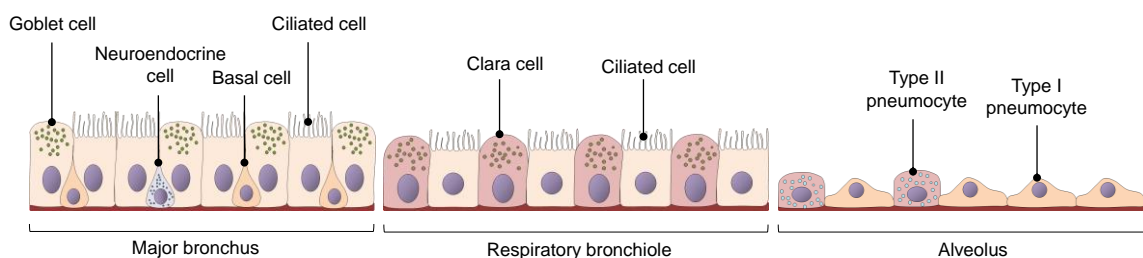


**Figure 1.1 – Cellular composition of human airway epithelia.**

## 1.2       Lung histopathology

Approximately 98% of primary lung tumours are carcinomas and are classified according to their histopathological characteristics (Travis *et al.*, 1995). The two major subtypes, which differ significantly with respect to treatment and prognosis, are small cell lung carcinoma (SCLC) (approximately 20%) and non-small cell lung carcinoma (approximately 80%). NSCLC may be further classified into adenocarcinoma (35 – 40%), squamous cell carcinoma (25 – 30%), large cell carcinoma (10 – 15%) and mixed subtypes (Califano *et al.*, 2012). Lung carcinomas present centrally (arising from major bronchi) or peripherally (arising from alveoli/respiratory bronchioles).

### 1.2.1       Small cell lung carcinoma

It is currently unclear whether SCLC arises from a preneoplastic lesion, although this tumour type is believed to originate predominantly from neuroendocrine cells within major bronchial epithelium (Park *et al.*, 2011; Sutherland *et al.*, 2011). SCLC is characterised by small cells (< 21 µm) with 'scant' cytoplasm and ill-defined borders (Fig 1.3 A). Characteristic nuclear defects include granular ('salt and pepper') chromatin, absent/undefined nucleoli and nuclear moulding. Consistent with its putative neuroendocrine origin, SCLC expresses several neuroendocrine markers, including neurosecretory vesicles, neural cell adhesion molecule, and neuropeptide growth factors (Franklin, 2000). Up to 30% of SCLCs are interrupted with discrete regions of non-small cell components, derived from all NSCLC subtypes. A tumour of this kind is classified as a combined SCLC (c-SCLC), although it is treated for standard SCLC (Wagner *et al.*, 2009). Rarely, SCLC can also be associated with less common lung malignancies, such as carcinosarcoma, giant cell carcinoma and spindle cell carcinoma (Summermann *et al.*, 1990; Fishback *et al.*, 1994; Tsubota *et al.*, 1992).

## 1.2.2 Non-small cell lung carcinoma (NSCLC)

### 1.2.2.1 Squamous cell carcinoma (SqCC)

SqCC develops from preneoplastic lesions within major bronchial epithelium. These lesions begin with basal cell hyperplasia, squamous metaplasia, dysplasia (mild, moderate, severe) and eventually progress to carcinoma *in situ* (CIS); a state that may persist for several years. Hyperplasia and squamous metaplasia alone are not regarded as preneoplastic changes, whereas dysplasia and CIS are *bona fide* preneoplastic lesions with a significant chance of progression (Jeremy George *et al.*, 2007). Mild dysplasia involves cell crowding and mild cytological atypia restricted to the basal layer (lower third of the epithelium). Moderate dysplasia involves cell crowding and vertically-orientated nuclei within the lower two-thirds of the epithelium, and mitoses are present in the lower third. Severe dysplasia involves greater cytological atypia, mitoses are present in the lower two-thirds of the epithelium, and there is virtually no maturation at the surface (Kerr, 2001). CIS is characterised by severe cytological atypia, cell crowding and scattered mitotic figures, which all occur throughout the entire epithelium (Lantuejoul *et al.*, 2009).

Histologically, SqCC comprises large, flattened cells with a high cytoplasm to nucleus ratio, and is characterised by intercellular bridging (desmosomes) and keratinisation, which may be in the form of scattered keratin pearls (concentric rings of squamous epithelium surrounding a core of acellular keratin) (Figure 1.3 B, centre) or intracytoplasmic keratin. These keratinised features are prominent in well-differentiated tumours and only focally present in poorly differentiated tumours (Churg *et al.*, 2005). Rare variants of SqCC include papillary, small cell, clear cell and basaloid, the latter of which gives poor prognosis (Travis *et al.*, 1999; Moro *et al.*, 1994).

## 1.2.2.2    Adenocarcinoma (AC)

Lung AC represents a group of heterogeneous tumours and, as such, histological classification is challenging. In light of recent advancements in our understanding of lung adenocarcinomas, a revised classification has been proposed by the International Association for the Study of Lung Cancer (IASLC), American Thoracic Society (ATS) and European Respiratory Society (ERS) (Travis *et al.*, 2011). Accordingly, adenocarcinomas can be grouped into preinvasive lesions, minimally invasive adenocarcinoma and invasive adenocarcinomas. If the former two types of adenocarcinoma are completely resected, a disease-specific survival rate of near 100% is expected (Travis *et al.*, 2012).

Atypical adenomatous hyperplasia (AAH) and adenocarcinoma *in situ* (AIS; formerly bronchioloalveolar carcinoma) are believed to be precursor lesions for invasive disease. AAH follows a non-invasive 'lepidic' growth pattern, in which tumour cells grow along and replace the surface of intact alveolar walls, without invading surrounding structures (Fig 1.2 A). Often found incidentally in resected lung tissue, AAH lesions have a maximum diameter of 5 mm and do not contain invasive foci. Histologically, cells resemble type II pneumocytes or Clara cells, exhibit mild atypia and frequently contain dense nuclear inclusions (Kitamura *et al.*, 1999; Mori *et al.*, 2001).

AIS is the next stage of a stepwise progression from pulmonary AC to invasive AC. This lesion also follows a lepidic growth pattern with no stromal invasion, but is larger than AAH, typically between 5 mm and 3 cm (Figure 1.2 B). Defining histological features of AIS include tall, columnar, closely packed tumour cells (giving a 'picket fence' pattern), cell crowding, cell stratification, coarse chromatin and prominent nuclei (Travis *et al.*, 2011).

AIS progresses to minimally invasive adenocarcinoma (MIA), which is distinct from AIS only by the presence of a ≤ 5 mm stromal invasion that follows a non-lepidic growth pattern. This lesion is not associated with vascular/pleural infiltration or necrosis (Borczuk *et al.*, 2009; Yim *et al.*, 2007; Maeshima *et al.*, 2010). AIS and MIA lesions are generally non-mucinous; only rare cases of mucinous lesions have been reported (Yoshizawa *et al.*, 2011).
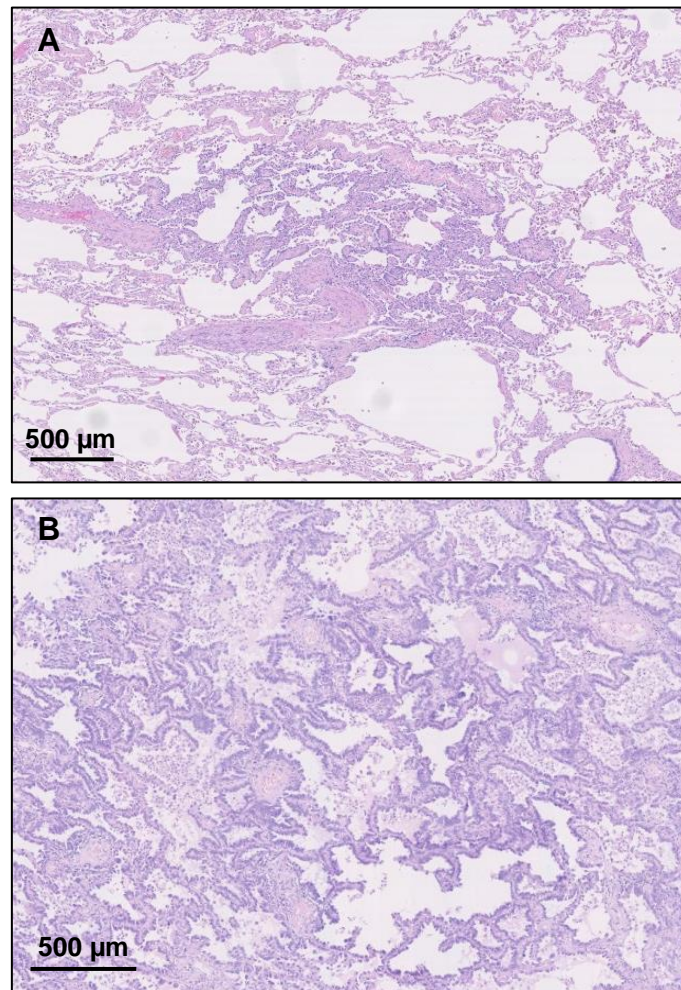
**Figure 1.2 – Preneoplastic stages of lung adenocarcinoma**. (A) Atypical adenomatous hyperplasia. (B) Adenocarcinoma *in situ*. Images kindly provided by Dr David Moore (histopathologist, Department of Cancer Studies, University of Leicester).

Invasive lung ACs may show up to five distinct patterns of growth. These are acinar, lepidic, papillary, micropapillary and solid with mucin production. The recent IASLC classification recommends that ACs showing multiple growth patterns, which account for up to 90% of cases, are identified by their predominant growth pattern (Travis *et al.*, 2011). Acinar growth presents histologically as irregular glandular structures of varying shapes and sizes, consisting of atypical epithelial cells with hyperchromatic nuclei and prominent nucleoli (Yoshizawa *et al.*, 2011). The malignant cells and/or glands themselves may or may not contain mucin (Travis *et al.*, 2013). As previously discussed in the context of pre-invasive lesions, lepidic growth is a non-invasive growth pattern involving mild-to-moderately atypical cells that proliferate along pre-existing alveolar walls, causing a slight thickening. A lepidic-predominant invasive AC is distinguished from AIS/MIA by the presence of an invasive focus measuring > 5 cm in its greatest diameter, which invades lymphatics, vasculature and/or pleura, or if the tumour contains necrotic lesions. The invasive focus, by definition, will comprise cells following an acinar, papillary, micropapillary or solid growth pattern (Travis *et al.*, 2013). Papillary growth is characterised by the presence of papillary structures, which are formed by the growth of tumour cells around fibrovascular cores. With micropapillary growth, these structures do not contain fibrovascular cores; instead, tumour cells form papillary 'tufts,' which may or may not be connected to the surrounding alveolar walls (Miyoshi *et al.*, 2003). Solid growth with mucin production comprises sheets of poorly differentiated polygonal cells with no evidence of other growth patterns. At least five cells should stain positively for intracytoplasmic mucin, per each of two high-power fields analysed. Solid predominant ACs may appear histologically similar to poorly differentiated SqCCs and large-cell carcinomas, which can also express intracytoplasmic mucin (Travis *et al.*, 2011). Rare growth patterns of invasive AC include foetal, colloid (mucinous), signet ring, clear cell and mucinous cystadenocarcinoma (Travis *et al.*, 2004).

### 1.2.2.3    Large cell carcinoma (LCC)

This term is assigned to poorly differentiated tumours that lack obvious small cell, squamous or glandular architecture. It is likely that LCC represents poorly differentiated squamous cell carcinomas and adenocarcinomas, and therefore the cell types of origin are likely to be multiple (Kitamura *et al.*, 2008). Defining features include sheets/nests of large, round to polygonal cells with pale-staining cytoplasm, vesicular nuclei and prominent nucleoli (Fig 1.3 D) (Travis *et al.*, 2004). Several uncommon variants include basaloid carcinoma, lymphoepithelioma-like carcinoma, clear cell carcinoma, LCC with rhabdoid phenotype and large cell neuroendocrine carcinoma. The latter rare variant belongs to a subset of neuroendocrine lung carcinomas (one of which is SCLC), and has a poor prognosis (Takei *et al.*, 2002).
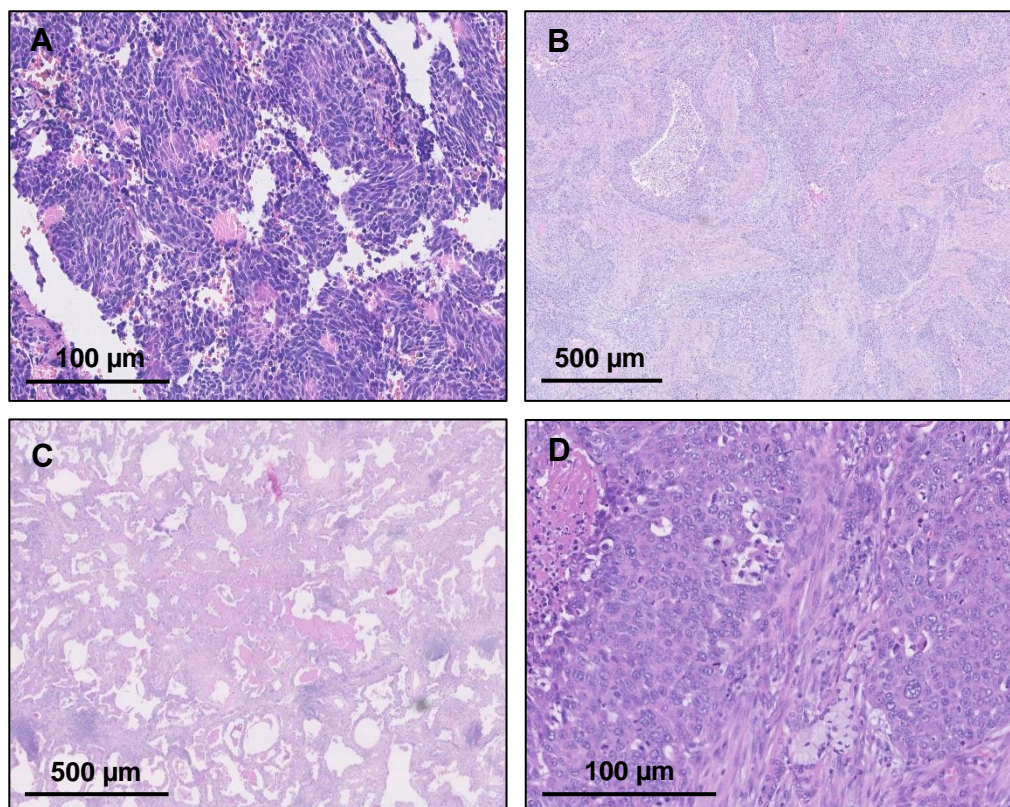


**Figure 1.3 – Histological appearance of lung carcinoma subtypes.** (A) Small cell lung carcinoma. (B) Squamous cell carcinoma. (C) Adenocarcinoma (lepidic predominant). (D) Large cell carcinoma. Images kindly provided by Dr David Moore (histopathologist, Department of Cancer Studies, University of Leicester).

## 1.3  Disease staging and prognosis

SCLC and NSCLC are staged by the TNM system, developed by the American Joint Committee on Cancer (AJCC) and the International Union Against Cancer. The most recent ($7^{th}$) edition adopts recommendations made by the International Association for the Study of Lung Cancer (IASLC) (Edge *et al.*, 2010). The TNM staging system combines three pieces of information: the size of the primary tumour (T), the extent of regional lymph node involvement (N) and the extent of metastasis to distant tissues (M) (Table 1.1). Once T, N and M scores are assigned, they are combined into an overall stage (Table 1.2). Although the $7^{th}$ edition of the TNM staging system includes SCLC, these tumours are commonly staged as either 'limited-stage' disease (LD) or 'extensive-stage' disease (ED), reflecting the tendency to metastasise early. LD is defined as a tumour confined to one hemithorax, with the possibility of regional lymph node extension. ED is defined as tumour extension beyond the regional lymph nodes (Stahel *et al.*, 1989).

The overall prognosis of lung cancer is extremely poor, because over two-thirds of patients present with evidence of local (stage IIIA) or distant (stage IIIB/IV) metastases that are refractory to curative treatment. In the UK, one-year and five-year survival rates for SCLC and NSCLC combined are approximately 30% and 8%, respectively (Office for National Statistics, 2011). Table 1.3 shows UK one-year and five-year survival rates (SCLC and NSCLC combined) by tumour stage.

The prognosis of SCLC is particularly poor, due to the rapid growth rate and onset of metastasis. The overall five-year survival rate for SCLC is approximately 5%, whilst for ED (TNM: $T_{any}$, $N_{any}$, M1a/b) it is < 1%. Median survival for ED without treatment is 2 – 4 months.

| T | Tumour |
|---|---|
| TX | Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy |
| T0 | No evidence of primary tumor |
| Tis | Carcinoma *in situ* |
| T1 | Tumor ≤3 cm in greatest dimension, surrounded by lung or visceral pleura, without bronchoscopic evidence of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus) |
| T2 | Tumor >3 cm but ≤7 cm in greatest dimension, or tumor with any of the following features (T2 tumors with these features are classified T2a if ≤5 cm): involves main bronchus, ≥2 cm distal to the carina; invades visceral pleura (PL1 or PL2); or is associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the entire lung |
| T3 | Tumor >7 cm or one that directly invades any of the following: parietal pleural (PL3) chest wall (including superior sulcus tumors), diaphragm, phrenic nerve, mediastinal pleura, or parietal pericardium or tumor in the main bronchus (<2 cm distal to the carinab but without involvement of the carina) or associated atelectasis or obstructive pneumonitis of the entire lung or separate tumor nodule(s) in the same lobe |
| T4 | Tumor of any size that invades any of the following: mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, carina, or separate tumor nodule(s) in a different ipsilateral lobe |
| N | Nodes |
| NX | Regional lymph nodes cannot be assessed |
| N0 | No regional lymph node metastasis |
| N1 | Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes, including involvement by direct extension |
| N2 | Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s) |
| N3 | Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s) |
| M | Metastasis |
| M0 | No distant metastasis |
| M1 | Distant metastasis |
| M1a | Separate tumor nodule(s) in a contralateral lobe tumor with pleural nodules or malignant pleural (or pericardial) effusion |
| M1b | Distant metastasis (in extrathoracic organs) |

**Table 1.1 – TNM staging of lung tumours (7th Edition).** Adapted from Edge *et al.*, 2010.

|  |  | T | N | M |
|---|---|---|---|---|
| **Occult carcinoma** |  | TX | N0 | M0 |
| **Stage 0 (AIS)** |  | Tis | N0 | M0 |
| **Stage I** | **A** | T1a, T1b | N0 | M0 |
|  | **B** | T2a | N0 | M0 |
| **Stage II** | **A** | T1a, T1b | N1 | M0 |
|  |  | T2a | N1 | M0 |
|  |  | T2b | N0 | M0 |
|  | **B** | T2b | N1 | M0 |
|  |  | T3 | N0 | M0 |
| **Stage III** | **A** | T1, T2 | N2 | M0 |
|  |  | T3 | N1, N2 | M0 |
|  |  | T4 | N0, M1 | M0 |
|  | **B** | T4 | N2 | M0 |
|  |  | Any T | N3 | M0 |
| **Stage IV** |  | Any T | Any N | M1a, M1b |

**Table 1.2 – TNM stage groupings.** Adapted from UyBico *et al.*, 2010.

| Stage | One-year survival (%) | Five-year survival (%) |
|---|---|---|
| I | 47% | 42% |
| II | 39% | 23% |
| III | 31% | 10% |
| IV | 23% | 2% |
| Unknown | 27% | 2% |

**Table 1.3 – NSCLC survival rates by tumour stage.** Data taken from Department of Health, 2011.

## 1.4       Screening and detection

Screening in medicine is the process of detecting a disease, or predicting the probability of development of a disease, in individuals showing no clinical signs and symptoms of the disease. The overall aim of screening is to reduce the incidence and mortality of a particular disease by early detection (Bretthauer and Kalager, 2013).

### 1.4.1       Imaging

To date, screening efforts for lung cancer have focused principally on chest radiography and computed tomography (CT). Radiographic screening studies have consistently failed to show a reduction in disease-specific mortality, and this was also the outcome of the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomised trial (Frost *et al.*, 1984; Flehinger *et al.*, 1984; Fontana *et al.*, 1991; Oken *et al.*, 2011). This is likely because small nodules are not detectable using radiography, and larger, benign nodules lead to false-positives (Dhillon *et al.*, 2013). The advent of low-dose CT imaging prompted the use of CT in the screening setting, because it was shown to detect small lung nodules at lower doses of radiation than conventional CT. Early single-arm studies consistently reported low-dose CT to be superior to chest radiography in the detection of small lung nodes, a significant portion of which were confirmed as malignant (Kaneko *et al.*, 1996; Sone *et al.*, 1998; Henschke *et al.*, 1999; Henschke *et al.*, 2001). Further studies recruiting current and former asymptomatic heavy-smokers reported the ability of low-dose CT to detect early-stage lung cancers (Swensen *et al.*, 2005; Diederich *et al.*, 2002; Pastorino *et al.*, 2003; Sobue *et al.*, 2002). Despite these encouraging results, a meta-analysis of six randomised trials assessing the utility of low-dose CT in lung cancer screening found a high rate of false-positive lesions (235 benign nodules per 1,000 individuals screened) (Gopal *et al.*, 2010). The largest randomised trial comparing low-dose CT to radiography in the screening of lung cancer was the National Lung Screening Trial (NLST), which recruited 53,454 current and former heavy smokers (National Lung Screening Trial Research Team *et al.*, 2011b). Individuals were screened at

baseline, year one, year two, and then monitored for around 6.5 years. The CT arm generated a positive rate of 27% at baseline, with 96% confirmed as false-positives. Overall, the percentage of stage 1 cancers identified through CT was 63%, compared to 47.6% identified through radiography. The overall number of disease-specific deaths throughout the trial was 354 in the CT arm and 442 in the radiography arm, which corresponds to a 20.3% decrease in disease-related deaths with CT screening (National Lung Screening Trial Research Team *et al.*, 2011a). The results from this trial prompted the National Comprehensive Cancer Network (NCCN), an alliance of twenty-one American cancer organisations, to publish recommendations for lung cancer screening in high-risk individuals using helical low-dose CT. NCCN defines high-risk individuals as either those aged 50-74 with a 30+ pack-year history of smoking tobacco, including those who have quite within the last 15 years, or those aged 50+ with a 20+ pack year history of tobacco smoking, and one additional risk factor (including cancer history, lung disease history, family history and carcinogen exposure) (Wood *et al.*, 2012). Such a screening programme has not yet been adopted in the UK.

### 1.4.2      Sputum cytology

Two early randomised lung cancer screening trials, namely the Johns Hopkins Lung Project and the Sloan-Kettering Lung Study, investigated the efficacy of annual chest radiography with and without sputum cytology at four-month intervals (Frost *et al.*, 1984; Melamed *et al.*, 1984). Both trials reported no significant benefit of sputum cytology, in terms of disease-specific mortality, but these trials failed to follow-up patients long-term and were not sufficiently powered to detect a significant benefit (Doria-Rose *et al.*, 2009). However, several combined analyses of these screening trials also report only a modest reduction in disease-specific mortality with sputum cytology (Rate Ratio [RR] 0.88; 95% confidence interval [CI], 0.74-1.05), and therefore this screening technique is not in clinical use (Doria-Rose *et al.*, 2009; Manser *et al.*, 2004).

## 1.5     Treatment and palliation

### 1.5.1     SCLC

The tendency of SCLC to metastasise early renders curative surgical resection impossible in most cases; only patients with a small, confined lesion and confirmed node-negativity (T1N0M0) may benefit from surgery (Yu *et al.*, 2010). SCLC is highly sensitive to chemotherapy and radiotherapy, and these treatments can demonstrate remarkable efficacy in shrinking the primary tumour.

Combined chemotherapy and radiotherapy (chemoradiotherapy) is the first-line treatment for LD SCLC. Patients that cannot tolerate concurrent chemoradiotherapy may be given sequential chemoradiotherapy if a response to chemotherapy is evident. The addition of radiotherapy has been shown to reduce intrathoracic recurrence by 25% in LD (Warde and Payne, 1992). The chemotherapy regimen for LD SCLC consists of a platinum-based agent (carboplatin/cisplatin) with a topoisomerase I/II inhibitor (irinotecan/etoposide), with cisplatin and etoposide as first-line (Pujol *et al.*, 2000; Sundstrom *et al.*, 2002). If a patient with LD SCLC responds adequately to radiotherapy, prophylactic cranial irradiation (PCI) can also be considered. This measure has been shown to increase three-year OS by 6% (Prophylactic Cranial Irradiation Overview Collaborative Group, 2000).

ED SCLC is treated with palliative intent, since distant metastases cannot be cured. The first-line treatment is chemotherapy, which is commonly a combination of cisplatin and etoposide (Jackman and Johnson, 2005). Radiotherapy may be given if there is evidence of a response to chemotherapy at distant metastatic sites.  Radiotherapy can also be used for palliation within symptomatic regions of disease (Lutz *et al.*, 2011). If a patient responds to first-line chemotherapy, PCI may be offered to reduce the incidence of brain metastases and lengthen both disease-free and overall survival (Slotman *et al.*, 2008).

Whilst SCLC typically responds very well to chemotherapy, approximately 80% of LD patients and virtually all ED patients eventually relapse (Schneider, 2008). If relapse is evident after 90 days of the completion of first-line treatment, the

patient is said to be sensitive to treatment. If relapse is evident within 90 days, the patient is said to be refractory to treatment (Hurwitz *et al.*, 2009). Around 25% of sensitive-relapse patients may benefit from second-line treatment, which is the topoisomerase I inhibitor, topotecan (von Pawel, 2003). However, second-line treatment only extends survival by several months. Refractory-relapse patients rarely respond to second-line treatment, and are thus more suited to palliative care (Ardizzoni *et al.*, 1997).

There are currently no targeted therapies approved for the treatment of SCLC. Many agents have been investigated in SCLC, including inhibitors of apoptosis (obatoclax), angiogenesis (bevacizumab, cediranib, marimastat, sorafenib, sunitinib, thalidomide, vandetanib), tyrosine kinases (gefitinib, imatinib), mTOR (everolimus, temsirolimus) and the insulin growth factor receptor. None of these agents have proved more beneficial for the treatment of SCLC over traditional chemotherapeutic agents (Lu *et al.*, 2013; Spigel, 2012). However, recent next-generation sequencing studies in SCLC have identified several actionable alterations that may be amenable to targeted therapy on an individual patient basis, such as mutations in *KRAS*, *EGFR*, *BRCA2* and *RICTOR* and possibly amplification of *MYC* (Ross *et al.*, 2014; Sos *et al.*, 2012).

### 1.5.2    NSCLC

In contrast to SCLC, NSCLC is relatively insensitive to chemotherapy and radiotherapy. Surgery is the standard radical treatment, but complete resection is only feasible in stages I/II, and occasionally in stage IIIA (Scott *et al.*, 2007). Surgery may be used for palliation in stage IV disease. Surgical resection may require the removal of an entire lung (pneumonectomy), a lobe (lobectomy), a small wedge-shaped section of a lobe (wedge resection) or a large section of a lobe (segmentectomy). During surgery, regional lymph nodes are removed and analysed for metastatic spread. In medically-fit patients, lobectomy is preferred over sublobar methods (wedge resection and segmentectomy); sublobar methods may be used for patients who cannot tolerate a lobectomy (Mery *et al.*, 2005). A sleeve lobectomy, in which a lobe and its supplying bronchial segment are removed, is indicated for the resection of more advanced or centrally-located tumours (Scott *et al.*, 2007).

Depending on the stage of disease, chemotherapy may be offered as a neoadjuvant therapy (stage IIIA), an adjuvant therapy (stage I-II), or without surgery (stage IIIB/1V or earlier stages where surgery is contraindicated) (Pignon *et al.*, 2008; NSCLC Meta-analyses Collaborative Group *et al.*, 2010). For advanced disease (stage IIIB/IV), chemotherapy is used as a first-line treatment, giving a modest improvement in overall survival at 1-year (NSCLC Meta-Analyses Collaborative Group, 2008). A typical regimen for NSCLC chemotherapy combines a platinum-based agent (cisplatin/carboplatin) with another drug such as a mitotic spindle poison (docetaxel, paclitaxel, vinorelbine), gemcitabine or pemetrexed (Chen *et al.*, 2010; Chang, 2011). Pemetrexed can be used for maintenance chemotherapy in patients with advanced non-squamous lung tumours who have not received the drug as first-line chemotherapy (Ciuleanu *et al.*, 2009). Radiotherapy is used to resect early-stage lung cancer when surgery is contraindicated by the presence of other health conditions, or when the tumour is inaccessible (Rowell and Williams, 2001). Radiotherapy is also used when surgery does not provide complete resection of the tumour. If stage IIIA lung cancer is not suitable for surgery, chemoradiotherapy is offered.

Surgery is not considered curative for stage IIIB disease and, therefore, chemoradiotherapy is first-line (Robinson *et al.*, 2007). Radiation can be given alone at this stage for palliation of specific symptoms. Stage IV lung cancer is also not resectable, and chemotherapy is used alone to extend survival. Radiotherapy and surgery can be offered for palliation (Lester *et al.*, 2006).

For advanced disease, it is sometimes beneficial to add a targeted agent to the chemotherapeutic regimen. Tumours harbouring activating mutations in the *EGFR* gene, predominantly L858R and short deletions within exon 19, may be targeted with reversible tyrosine kinase inhibitors (TKIs), namely erlotinib and gefitinib. These agents may extend PFS by several months in *EGFR*-mutant patients, but their efficacy in *EGFR*-negative patients remains unclear (Paz-Ares *et al.*, 2010; Laurie and Goss, 2013). The *EML4-ALK* fusion, which is present in approximately 4% of NSCLCs, permits the use of the ALK/MET/ROS1 inhibitor, crizotinib. Like the EGFR inhibitors, crizotinib can extend PFS by several months (Forde and Rudin, 2012). Whilst targeted agents are associated with fewer side effects than standard chemotherapeutic agents, their efficacy is limited by the emergence of resistance mechanisms (Kobayashi *et al.*, 2005; Choi *et al.*, 2010). Resistance to EGFR TKIs is commonly acquired through the T790M mutation at the 'gatekeeper' residue within the ATP-binding pocket. Several studies have showed that this mutation is a subclonal event that occurs prior to initiation of EGFR TKI therapy (Wheeler *et al.*, 2010). Less commonly, acquired resistance can be caused by amplification of *MET* or mutation of genes downstream of EGFR (Stewart *et al.*, 2015). Resistance to ALK TKIs arises via secondary mutations in the tyrosine kinase domain of ALK or EGFR in approximately 60% of cases. Genomic amplification of *ALK* fusions is responsible for a further 15% of cases. Activation of *KIT*, *KRAS* and *IGFR1* is also implicated in acquired resistance, but their relative contributions are currently unclear (Isozaki *et al.*, 2015).

## 1.6     Genetic alterations in lung cancer and preneoplasia

### 1.6.1    NSCLC

Genomic sequencing efforts in NSCLC have revealed a complex and heterogeneous landscape of alterations, some of which appear to discriminate between molecular subtypes and smoking status. In addition, these efforts have led to the discovery of novel drivers that may be actionable with existing therapies or therapies currently undergoing clinical testing (Araujo *et al.*, 2015). A summary of large-scale genome sequencing studies in NSCLC is provided in Table 1.4, showing the most commonly altered genes.

The mutational burden of NSCLC is ten-fold higher in smokers than in never-smokers, with a positive correlation between smoking exposure and mutation number (Govindan *et al.*, 2012a). Moreover, C>A transitions are particularly abundant among smokers, whereas C>T transitions are more abundant among never-smokers (Govindan *et al.*, 2012a; Ding *et al.*, 2008a; Lee *et al.*, 2010; Imielinski *et al.*, 2012). Interestingly, smoking is associated with alterations in specific gene sets; *KRAS*, *TP53*, *BRAF, STK11* and JAK2/JAK3 are commonly mutated in smokers, whereas *EGFR* and *PTEN* mutations and *ALK*, *ROS1* and *RET* fusions are common in never-smokers (Govindan *et al.*, 2012a; Seo *et al.*, 2012; Shigematsu *et al.*, 2005; An *et al.*, 2012).

The *KRAS* gene is mutated in approximately 25% of NSCLC. The majority of mutations are missense, affecting codons 12, 13 and 61, resulting in hyperactivation of *KRAS* signalling (Johnson *et al.*, 2012). *EGFR* mutations are found in 15 – 40% of NSCLC, with a high prevalence among East Asians (Yoshida *et al.*, 2007; Cortes-Funes *et al.*, 2005; Tokumo *et al.*, 2005). Up to 90% of *EGFR* mutations are the L858R missense mutation (exon 21) and small, in-frame exon 19 deletions, and these are currently screened for in routine practice since they confer sensitivity to TKIs (Ladanyi and Pao, 2008). *KRAS* and *EGFR* mutations are mutually exclusive and have been identified in both preneoplastic lung lesions (Chapter 1.6.3) and founder clones of overt NSCLC, strongly suggesting these genes are early drivers (Govindan *et al.*, 2012a; Yoshida *et al.*, 2005; Tang *et al.*, 2005; Marchetti *et al.*, 2011). *TP53* is mutationally inactivated

in 40 – 60% of NSCLC, independent of *KRAS* and *EGFR* status (Mogi and Kuwano, 2011; Mao, 2001). Mutations are predominantly missense, clustering in exons 5 – 8 (DNA-binding domain) (Vousden and Lu, 2002; Yokota and Kohno, 2004; Mechanic *et al.*, 2005). Several other genes are reported to be mutated at varying frequencies, including *BRAF*, *ERBB2*, *FGFR2*, *PIK3CA*, *PTEN* and *RB1* (Table 1.4).

Somatic copy number alteration (SCNA) of key genes is also a major driver in NSCLC. Amplification of *MET* is observed in 2 - 4% of TKI-naïve patients and rises to around 20% in patients receiving TKI therapy as a mechanism of drug resistance (Go *et al.*, 2010; Bean *et al.*, 2007; Chen *et al.*, 2009; Cappuzzo *et al.*, 2009a; Cappuzzo *et al.*, 2009b). The frequency of *EGFR* amplification has been reported at 4 – 14% and copy number is closely correlated with *EGFR* mutation status (Chang *et al.*, 2008; Liang *et al.*, 2010). Although controversial, high-level amplification of *EGFR* has been shown to confer sensitivity to TKIs (Cappuzzo *et al.*, 2005; Hirsch *et al.*, 2005; Sone *et al.*, 2007). Several SCNAs are strongly smoking associated, such as amplifications in *FGFR1*, *PIK3CA* and *SOX2*. These alterations predominantly occur in SqCC at frequencies of 5 – 8%, around 33% and 20 – 72%, respectively (Kawano *et al.*, 2006; Yamamoto *et al.*, 2008; Wang *et al.*, 2012; Wang *et al.*, 2014a; Yuan *et al.*, 2010; Wilbertz *et al.*, 2011; Brcic *et al.*, 2012; Bass *et al.*, 2009; Hussenet *et al.*, 2010a). Amplification of *ERBB2* is uncommon (around 1% of ACs) but has been proposed as a mechanism of resistance to TKI therapy (Takezawa *et al.*, 2012).

A subset of NSCLCs are driven by gene fusions. The most common gene fusion is *EML4-ALK* (3 – 6%), which is generated by the translocation of an inverted region (variable composition) of *EML4* to exons 20 – 29 of *ALK* (Soda *et al.*, 2007). 1 – 2% of ACs are driven by *RET* fusions with *CCDC6*, *KIF5B* or *TRIM33* and all fusions exhibit sensitivity to TKI therapy (Kohno *et al.*, 2012; Lipson *et al.*, 2012).  A further 1 – 2% of ACs are driven by ROS1 fusions with diverse partners, including *CD74*, *EZR*, *FIG*, *LRIG*, *SCF34*, *SDC4*, *SLC34A2* and *TPM3* (Bergethon *et al.*, 2012; Yoshida *et al.*, 2013; Davies *et al.*, 2012; Takeuchi *et al.*, 2012; Rimkunas *et al.*, 2012). Interestingly, all *ALK*, *RET* and *ROS* fusions are

19

associated with light/never smokers, AC histology and are mutually exclusive of other drivers (Thunnissen *et al.*, 2012).

| Study *(et al)* | Year | NSCLC subtype (*n*) | Commonly altered genes (top five) | Reference |
|---|---|---|---|---|
| Lin | 2016 | AC (61) | *EGFR, KRAS, PIK3CA, BRAF, EML4-ALK* (fusion) | Shao *et al.*, 2016 |
| Deeb | 2015 | Not specified (13) | *KRAS, ERBB2, EGFR, PTEN, -* | Deeb *et al.*, 2015 |
| Qiu | 2015 | AC (89) | *TP53, EGFR, KRAS, PIK3CA, BRAF* | Qiu *et al.*, 2015 |
| Hagemann | 2015 | AC (147), SqCC (36), LCC (4), other (15) | *KRAS, EGFR, PIK3CA, BRAF, PTEN* | Hagemann *et al.*, 2015 |
| TCGA Research Network | 2014 | AC (230) | *TP53, KRAS, KEAP1, STK11, EGFR* | Cancer Genome Atlas Research Network, 2014 |
| Govindan | 2012 | AC (16), LCC (4) | *KRAS, LRP1B, TP53, MET, EGFR* | Govindan *et al.*, 2012b |
| Imielinski | 2012 | AC (183) | *TP53, KRAS, EGFR STK11, KEAP1* | Imielinski *et al.*, 2012 |
| Lipson | 2012 | Not specified (24) | *KRAS, TP53, STK11, LRP1B, JAK2* | Lipson *et al.*, 2012 |
| TCGA Research Network | 2012 | SqCC (178) | *TP53, CDKN2A, PTEN, PIK3CA, KEAP1* | Cancer Genome Atlas Research Network, 2012 |
| Xiong | 2012 | AC (3), SqCC (5), other (6) | *TP53, MXRA5, EGFR, PIK3CA, KRAS* | Xiong *et al.*, 2012 |
| Ding | 2008 | AC (188) | *TP53, KRAS, STK11, EGFR, LRP1B* | Ding *et al.*, 2008b |

**Table 1.4 – Commonly altered genes in NSCLC as determined by large-scale genomic analyses.** AC, adenocarcinoma; LCC, large cell carcinoma; SqCC, squamous cell carcinoma.

## 1.6.2     SCLC

The genomic landscape of SCLC is less well characterised. Genomic characterisation of this tumour subtype is challenging due to the high mutational burden and the difficulty in obtaining clinical material for research. This has led to cell lines becoming a vital resource for genomic analyses (Semenova *et al.*, 2015). It is known, however, that SCLC exhibits a very high mutational burden (second to colorectal cancer), consistent with tobacco carcinogenesis, and is frequently driven by copy number alterations and inactivating mutations in tumour suppressors such as *TP53*, *RB1* and *PTEN* (Toyooka *et al.*, 2003; Wistuba *et al.*, 2000a) Activating mutations in oncogenes such as *EGFR*, *PIK3CA* and *KRAS* are relatively rare.

The most common genetic alterations in SCLC are mutational inactivation/deletion of *TP53* and *RB1*, which are observed in 66 – 95% of cases (Wistuba *et al.*, 2001; Arriola *et al.*, 2008; Mori *et al.*, 1990). Compared with other lung cancer subtypes, SCLC has the highest rate of allelic loss within adjacent 'normal' bronchial epithelium. Interestingly, 17p13 deletions (*TP53*) are particularly frequent, suggesting *TP53* loss may be an initiating event in the development of these tumours (Wistuba *et al.*, 2000b). Mutational inactivation of *PTEN* is observed in 4 – 16% of cases and 20 – 40% of SCLCs harbour mutations and/or amplifications in other genes of the PI3K pathway, leading to pathway hyperactivation (Ross *et al.*, 2014; Yokomizo *et al.*, 1998; Peifer *et al.*, 2012; Umemura *et al.*, 2014).

Many copy number alterations in SCLC are also observed in NSCLC. Loss of large segments on chromosome 3p is common to all lung cancer subtypes and suggests key TSGs are likely contained within this region (Otterson *et al.*, 1992). However, few TSGs within 3p have been identified to date. In SCLC, a pattern of arm-level LOH is frequently observed in chromosomes 4p, 4p, 4q, 10q, 13q (*RB1*), 1q and 17p (*TP53*), and appears to be unique to this tumour subtype (Ried *et al.*, 1994; Virmani *et al.*, 1998; Shivapurkar *et al.*, 1999; Staaf *et al.*, 2013). Genomic amplifications of the *MYC* family genes (*MYC*, *MYCL1* and *MYCN*) are well documented in SCLC and are mutually exclusive events (Sos *et al.*, 2012).

Reported amplification frequencies vary widely, but a previous array comparative genomic hybridisation study identified *MYC* family amplification in 81.8% of SCLC tumours (n = 33) and 61.5% of SCLC cell lines (n = 13) (Voortman *et al.*, 2010). Interestingly, *SOX2* amplification, which is highly prevalent in SqCC, has also been identified in SCLC by two recent comprehensive genomic analyses. However, the reported frequencies vary (4%, n = 98; 27%, n = 53) (Ross *et al.*, 2014; Rudin *et al.*, 2012). Another recently implicated gene is *FGFR1*, which is amplified in 4 – 8% of cases (Peifer *et al.*, 2012; Schultheis *et al.*, 2013; Thomas *et al.*, 2014; Zhang *et al.*, 2015)

Recent large-scale next-generation sequencing analyses by Rudin *et al.*, (2012), Peifer *et al.*, (2012), Ross *et al.*, (2014), George *et al.*, (2015) and Umemura *et al.*, (2015) have highlighted several key drivers of the disease, some of which may be actionable with current targeted therapies. A summary of genetic drivers in SCLC is provided in Table 1.5.

| Gene | Alteration | Frequency (%) | *n* | Study |
|---|---|---|---|---|
| *FGFR1* | A | 6 | 63 | Peifer *et al.*, 2012 |
| | A | 6 | 251 | Schultheis *et al.*, 2013 |
| | A | 8 | 68 | Thomas *et al.*, 2014 |
| | A | 4 | 98 | Ross *et al.*, 2014 |
| | A | 8 | 77 | Zhang *et al.*, 2015 |
| *MYCL1* | A | 10 | 183 | Johnson *et al.*, 1992 |
| | A | 8 | 63 | Sos *et al.*, 2012 |
| | A | 8 | 63 | Peifer *et al.*, 2012 |
| *MYCN* | A | 7 | 183 | Johnson *et al.*, 1992 |
| | A | 5 | 63 | Sos *et al.*, 2012 |
| | A | 6 | 63 | Peifer *et al.*, 2012 |
| *PTEN* | IM | 16 | 44 | Yokomizo *et al.*, 1998 |
| | IM | 10 | 29 | Peifer *et al.*, 2012 |
| | IM | 4 | 51 | Umemura *et al.*, 2014 |
| | IM | 5 | 98 | Ross *et al.*, 2014 |
| *RB1* | IM | 66 | 29 | Peifer *et al.*, 2012 |
| | D | 76 | 29 | Peifer *et al.*, 2012 |
| | IM | 80 | 110 | George *et al.*, 2015 |
| | D | 95 | 110 | George *et al.*, 2015 |
| | R | 14 | 110 | George *et al.*, 2015 |
| *SOX2* | A | 4 | 98 | Ross *et al.*, 2014 |
| | A | 27 | 56 | Rudin *et al.*, 2012 |
| *TP53* | IM | 93 | 29 | Peifer *et al.*, 2012 |
| | D | 66 | 29 | Peifer *et al.*, 2012 |
| | IM | 94 | 110 | George *et al.*, 2015 |
| | D | 95 | 110 | George *et al.*, 2015 |
| | T | 5 | 110 | George *et al.*, 2015 |

**Table 1.5 – Frequency of key genetic drivers in SCLC.** D, deletion (loss); A, genomic amplification; IM, inactivating mutation; T, translocation.

### 1.6.3      Lung preneoplasia

Genetic alterations that promote the development of preneoplasia are a fundamental consideration in the development of strategies for early diagnosis and treatment. Relatively few studies have investigated the genetic changes associated with lung preneoplasia, reflecting the difficulty in identifying individuals with specific lesions of interest. Rather, these lesions are often incidental biopsy findings in the lungs of patients with established cancer.

Allelic loss at 3p is one of the earliest genetic alterations in lung tumourigenesis, occurring in around 78% of bronchial preneoplastic lesions and 96% of lung tumours. A particularly interesting gene encoded within chromosome 3p is *FHIT*. This gene is subject to allelic loss in the majority of SCLCs and NSCLCs and has been confirmed as a *bona fide* TSG in lung cancer through a series of *in vitro* and *in vivo* studies (Sozzi *et al.*, 1996; Antczak *et al.*, 2013; Chen *et al.*, 2013). In patients with established lung cancer, preneoplastic lesions also exhibit aneusomy at chromosomes 7p12 and 8q24, which encompass the sequences for *EGFR* and *MYC*, respectively (Jonsson *et al.*, 2008).

Altered expression of p53 has been reported in lung preneoplasia, with around 25% of bronchial mild dysplastic lesions, 75% of CIS lesions and up to 60% of AAH lesions exhibiting increased protein expression (Brambilla *et al.*, 1998; Tominaga *et al.*, 2003). A possible explanation is that inactivating mutations in *TP53* extend the half-life of the protein and promote nuclear accumulation (Mattioni *et al.*, 2013). p53 accumulation in bronchial preneoplasia is associated with a high probability of invasive progression (Brambilla *et al.*, 1998; Ponticiello *et al.*, 2000; Jeanmart *et al.*, 2003). Thus, *TP53* inactivation may be a very early event in lung tumourigenesis.

Inactivation of *RB1* is not a feature of bronchial preneoplasia, but the retinoblastoma phosphorylation pathway is altered by LOH at 9p21, which encompasses *CDKN2A*. This implicates *CDKN2A* loss in early tumourigenesis of bronchial epithelium (Wistuba *et al.*, 1999; Chyczewski *et al.*, 2001).

Activating mutations in *EGFR* are common in peripheral preneoplasia, and are correlated with the severity of the lesion; around 3% of AAH lesions exhibit

mutations, compared to 10% of AIS lesions and 42% of overt adenocarcinomas (Yoshida *et al.*, 2005). Interestingly, patients with *EGFR*-mutant AC exhibit identical *EGFR* mutations in histologically 'normal' respiratory epithelium within and adjacent to the tumour. This phenomenon is not reported in *EGFR*-WT AC, suggesting *EGFR* mutations may be subject to a localised field effect in the respiratory epithelia, and implicating *EGFR* in early lung tumourigenesis (Tang *et al.*, 2005). *EGFR* mutations are uncommon in bronchial preneoplasia.

Oncogenic *KRAS* is observed in preneoplastic alveolar lesions, but decreases in frequency as the severity of the lesion increases. Sakamoto *et al* reported mutation frequencies of 33% of AAH, 12% of AIS, 8% of MIA and 0% of low-grade invasive ACs (Sakamoto *et al.*, 2007). Marchetti *et al* reported mutation frequencies of 36% in AIS and 22% of invasive ACs (Marchetti *et al.*, 2011). Therefore, *KRAS* mutations occur early in lung tumourigenesis, but the majority of mutant lesions do not progress to malignancy. This is also consistent with conditional mouse models of *Kras*^G12D^-induced lung cancer, whereby the majority of benign nodules do not progress to AC (Collado *et al.*, 2005). *KRAS* mutations may not be present in bronchial preneoplasia (Sugio *et al.*, 1994).

Amplification of 3q is thought to be the most common genetic alteration in SqCC. Consistently, *PIK3CA* is located within this region, and is frequently amplified in severely dysplastic bronchial lesions (Massion *et al.*, 2004). Another oncogene within 3p is *SOX2*, which has been identified as a possible driver of SqCC, and is also amplified in such lesions (McCaughan *et al.*, 2010; Schneider *et al.*, 2013).

## 1.7     Genetically engineered mouse models of lung preneoplasia

In humans, it is often difficult to ascertain whether benign lesions will progress to malignancy, since lesions are not trackable over time. Genetically engineered mouse (GEM) models of cancer have the advantage that lesions can be induced *de novo* through the activation of one or more genetic alterations. Mice can be monitored *in vivo* and harvested at defined time points to develop an understanding of tumourigenesis from initiation to metastasis. GEM models of preneoplasia have particular utility for the study of early neoplasia where either an exploratory approach is taken to identify early disease biomarkers, or where known genetic alterations in the preneoplastic lesions are evaluated in blood. Two well characterised GEM models that produce a spectrum of preneoplastic lung lesions are Cre/Lox-regulated models driven by $Braf^{V600E}$ and $Kras^{G12D}$.

### 1.7.1     Braf$^{+/LSL-V600E}$

The first GEM model of $Braf^{V600E}$-driven lung cancer was generated in the McMahon lab (Dankort *et al.*, 2007). In this model (Braf$^{+/LSL-V600E}$), a Cre/Lox-regulated mini-gene containing exons 15-18 is targeted downstream of exon 14 of *Braf*, such that in the absence of Cre recombinase, $Braf^{WT}$ is expressed at endogenous levels. In the presence of Cre recombinase, the mini-gene is excised and $Braf^{V600E}$ is expressed. Lung-specific expression is achieved through nasal administration of adenoviral Cre recombinase, the titre of which influences tumour multiplicity. Within 2 weeks, mice develop multiple hyperplastic regions of bronchial epithelium, progressing to papillary adenomas by 6-8 weeks of age. These lesions rarely progress to malignancy and, instead, show markers of senescence (Dankort *et al.*, 2007). Malignant progression can be achieved through concomitant inactivation of *Trp53* or *Ink4a/Arf*. A similar model of $Braf^{V600E}$-driven cancer was developed by the Pritchard lab in 2005 and can be induced by nasal delivery of adenoviral Cre recombinase as in the McMahon model (Mercer *et al.*, 2005). Figure 1.4 shows the preneoplastic features of the lung of a Braf$^{+/LSL-V600E}$ 12 weeks post-administration of Cre recombinase.
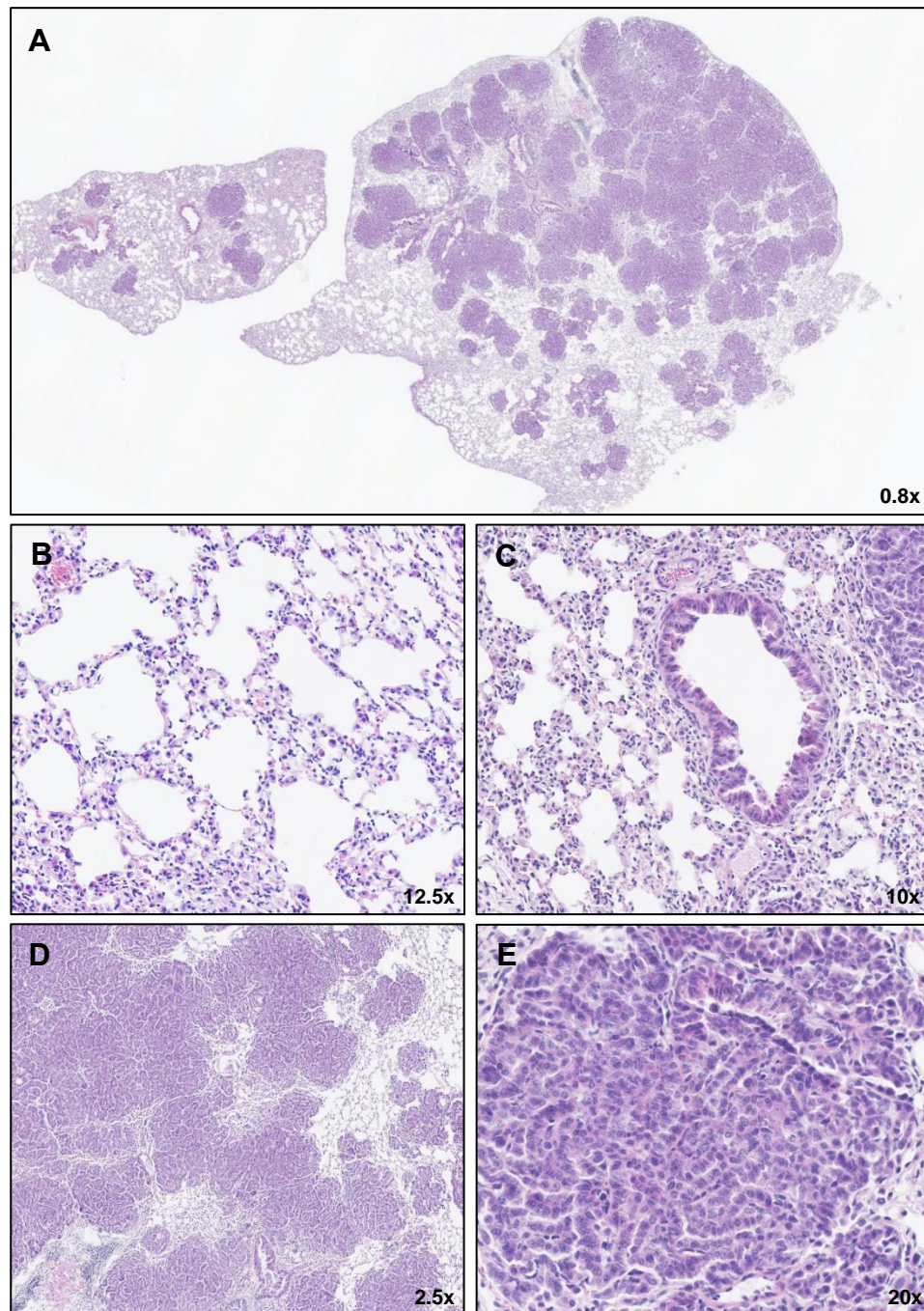
**Figure 1.4 – Histological appearance of lung lesions in the Braf$^{+/LSL-V600E}$ mouse.** (A) At 12 weeks post-infection with adenoviral Cre recombinase at 5 x 10$^7$ PFU, lungs exhibit widespread alveolar hyperplasia and numerous well-circumscribed adenomas. (B) Alveolar hyperplasia that resembles atypical adenomatous hyperplasia (AAH) observed in the human lung. (C) Bronchial epithelial hyperplasia with surrounding alveolar hyperplasia. (D) Multiple well-circumscribed adenomas. (E) High magnification of an adenoma in (D), showing a papillary predominant architecture. Adenocarcinomas are not evident at 12 weeks post-infection.

### 1.7.2        Kras<sup>+/LSL-G12D</sup>

The Kras$^{+/LSL\text{-}G12D}$ GEM model, developed by Tyler Jacks and colleagues (MIT), is a well characterised and extensively studied model that recapitulates the cognate human disease from the preneoplastic stages through to metastasis (Fig 1.5) (Jackson *et al.*, 2001). The model is regulated by LoxP-flanked transcriptional silencing (STOP) element upstream of exon 1, such that upon delivery of Cre recombinase, the STOP element is excised and *Kras*$^{G12D}$ is expressed at endogenous levels. In the absence of Cre recombinase, transcription is abrogated within the STOP sequence and no protein is produced. Like the Braf$^{+/LSL\text{-}V600E}$ model, spatiotemporal expression of *Kras*$^{G12D}$ in the lung is achieved by infecting mice intra-nasally with a recombinant adenovirus or lentivirus expressing Cre recombinase. Adenoviral administration gives moderate control over the tumour multiplicity and, more recently, the cell-of-origin through the delivery of adenoviral Cre strains that carry cell type-specific promoters (Sutherland *et al.*, 2011; Tao *et al.*, 2014).

In this thesis, the Pritchard (Braf$^{+/LSL\text{-}G12D}$) and Jacks (Kras$^{+/LSL\text{-}G12D}$) GEM models of early lung cancer development were investigated as models of circulating-free DNA in pre-malignant lung disease (Chapters 3 and 4).
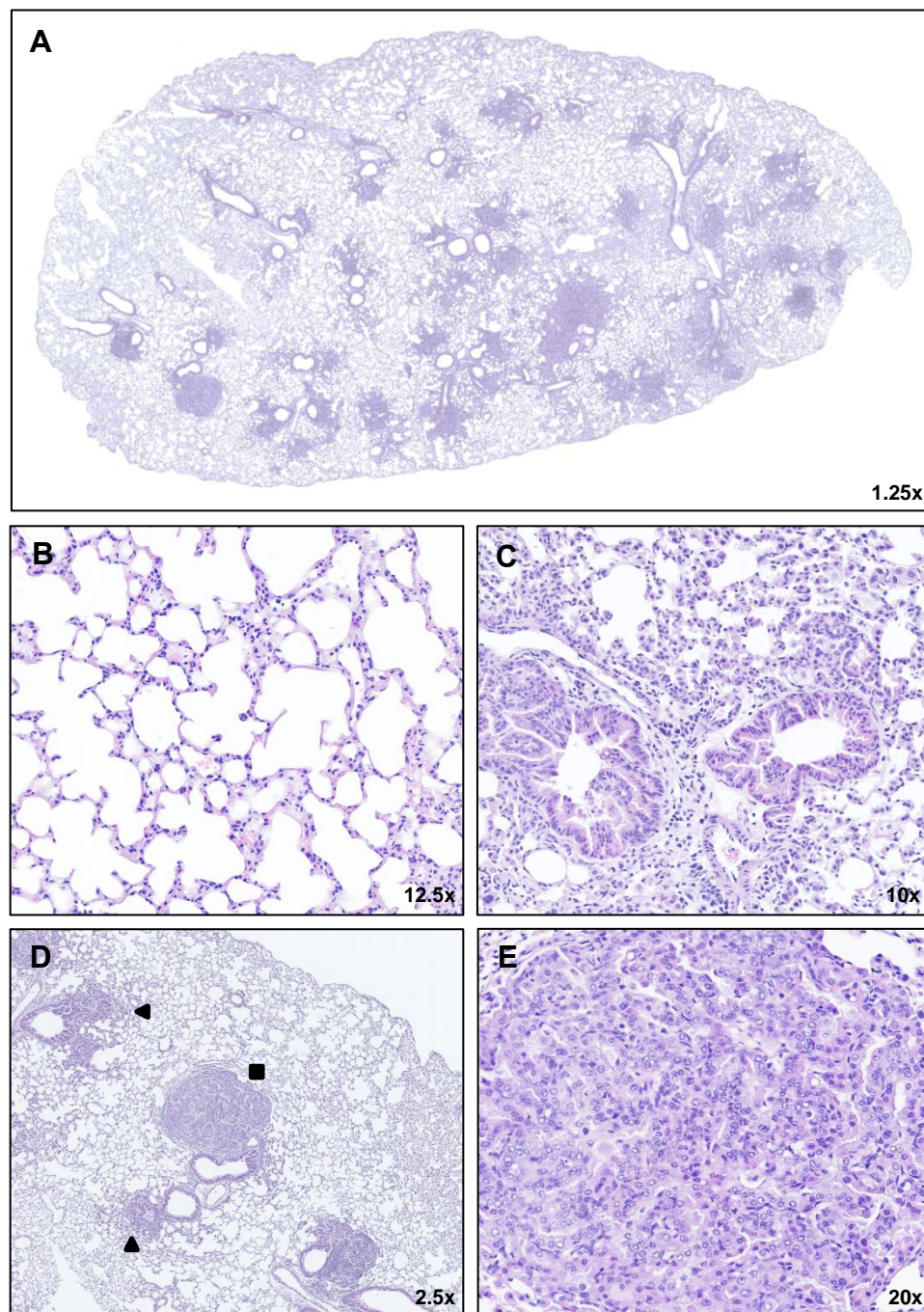
**Figure 1.5 – Histological appearance of lung lesions in the Kras$^{+/LSL\text{-}G12D}$ mouse.** (A) At 12 weeks post-infection with adenoviral Cre recombinase at 5 x 10$^7$ PFU, lungs exhibit widespread alveolar hyperplasia, multiple adenomas with adjacent bronchiolar involvement and bronchial epithelial hyperplasia. (B) Alveolar hyperplasia resembles atypical adenomatous hyperplasia (AAH) observed in the human lung. (C) Bronchial epithelial hyperplasia with surrounding alveolar hyperplasia. (D) Multiple adenomas with adjacent bronchiolar involvement. Adenomas may be well-circumscribed (■) or with lepidic borders (▲), also evident in (A). (E) High magnification of the central adenoma in (D), showing a papillary predominant architecture. Adenocarcinomas are not evident at 12 weeks post-infection.

## 1.8        Circulating biomarkers

Solid tumours are diagnosed and treated following histopathological analysis of tumour tissue. However, the procedure for obtaining tumour tissue is invasive and labour-intensive. Moreover, a tissue biopsy may be subject to spatial selection bias and fail to represent the histological and genetic heterogeneity of a tumour, potentially resulting in misguided treatment decisions. A 'liquid biopsy,' through circulating biomarkers, represents an alternative approach, since it is minimally invasive, repeatable, and may provide real-time snapshots of tumour dynamics (Table 1.6 and Fig 1.6)

Circulating biomarkers of interest in the management of cancer include circulating free DNA (cfDNA), circulating tumour cells (CTCs), small RNA (mRNA, miRNA), exosomes, microvesicles, platelets and plasma proteins (Gold *et al.*, 2015). However, none of these biomarker sources have been validated for clinical use in lung cancer (Xiang *et al.*, 2013).

|                             | **Tissue Biopsy**          | **Liquid Biopsy**   |
|-----------------------------|----------------------------|---------------------|
| **Invasiveness**            | Invasive                   | Minimally invasive  |
| **Accessibility**           | Dependent on tumour location | Always accessible |
| **Risk of complications**   | Post-operative risk        | Negligible risk     |
| **Length of process**       | Long and labour-intensive  | Fast and efficient  |
| **Longitudinal sampling**   | Not always feasible        | Feasible            |
| **Representative of tumour(s)** | No; spatial selection bias | Yes (in principal) |

**Table 1.6 – Tissue biopsy vs liquid biopsy.**
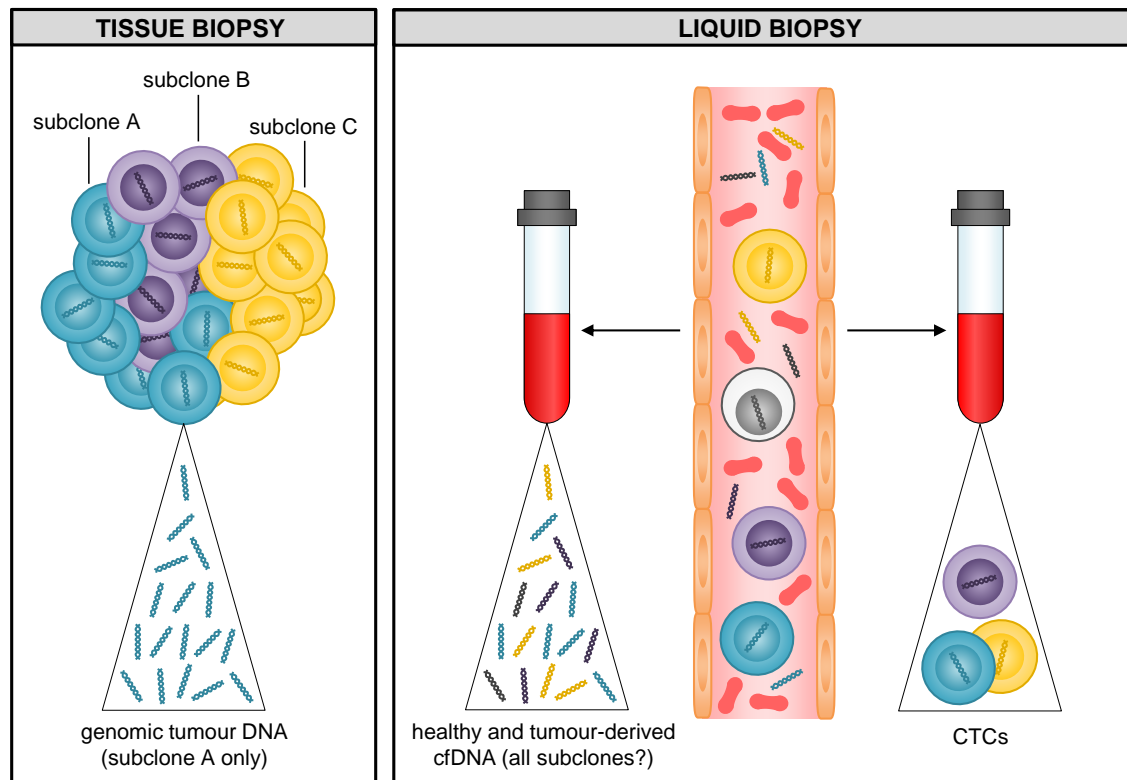
**Figure 1.6 – Blood based 'liquid biopsy.'** A tissue biopsy is invasive, subject to spatial selection bias and may not faithfully recapitulate genetic heterogeneity. cfDNA can be isolated from a blood sample and may represent the subclonal populations of a tumour. CTCs are under investigation as biomarkers for metastatic potential. Adapted from Figure 1, Fleischhacker & Schmidt, 2011.

### 1.8.1     Circulating-free DNA

### 1.8.1.1     Introduction

Nucleic acids (later confirmed as DNA) were first reported in the circulation over 60 years ago by Mandel and Métais (Mandel, P. and Métais, P., 1948). The discovery generated little interest until 1977, when cfDNA levels were first shown to be elevated in patients with cancer (Leon *et al.*, 1977). The tumour-derived component of cfDNA, termed circulating tumour DNA (ctDNA), was identified in 1994 by two independent studies in which *RAS* mutations were identified in the plasma of patients with pancreatic cancer and acute myelogenous leukemia respectively (Sorenson *et al.*, 1994; Vasioukhin *et al.*, 1994). These findings generated new enthusiasm to explore the clinical potential of circulating nucleic acids in patients with cancer.

There are several different mechanisms of cfDNA release. cfDNA fragments are typically 70 – 200 bp in length, and characteristic of apoptotic cleavage. In some patients with cancer, high-molecular weight DNA is also evident likely derived from necrotic cells (Giacona *et al.*, 1998; Jahr *et al.*, 2001; Jiang *et al.*, 2015). Basal cfDNA levels arise due to turnover of normal haematopoietic cells (Giacona *et al.*, 1998). Whilst cfDNA levels are generally elevated in cancer, particularly in advanced disease, this phenomenon is also recognised as a hallmark of other conditions including diabetes, systemic lupus erythematosus, sepsis and myocardial infarction, thus, increased cfDNA levels are not specific to cancer (Rumore and Steinman, 1990; Rhodes *et al.*, 2006; Rainer *et al.*, 2003). This realisation prompted the analysis of specific genetic and epigenetic alterations in cfDNA that mirror those of the solid tumour.

## 1.8.1.2    Biological importance of cfDNA

Intriguing *in vitro* and *in vivo* studies show that cfDNA from patients with cancer can potentially transform susceptible and seed distant metastases; the 'genometastasis' hypothesis (Garcia-Olmo *et al.*, 2000; Garcia-Olmo *et al.*, 2004). For example, cfDNA from tumour-bearing rats can transfect cancer cells in culture, and lung extracts from healthy rats contain tumour-specific markers after inoculation with plasma from tumour-bearing rats several weeks prior (Garcia-Olmo *et al.*, 1999b; Garcia-Olmo *et al.*, 1999a) and healthy murine cells cultured with cfDNA from tumour-bearing animals form xenografts when injected into immunodeficient mice. Lastly, 3 of 4 plasmas with mutant *KRAS* from patients with resected *KRAS*-mutated colorectal cancer were able to transform NIH-3T3 cells, which were subsequently shown to induce carcinomas when injected into nude mice (Garcia-Olmo *et al.*, 2010). These studies suggest that primary tumour resection does not necessarily eradicate circulating oncogenic DNA, and in some cases this DNA has transforming potential (Garcia-Olmo *et al.*, 2012).

Susceptible cells may also undergo oncogenic transformation via phagocytosis of tumour-derived apoptotic bodies. Bergsmedh *et al.* (2001) reported the ability of p53-deficient mouse embryonic fibroblasts (MEFs) to phagocytose apoptotic bodies derived from tumour cells, which induced loss of contact inhibition and anchorage independence *in vitro* and tumour formation *in vivo*.

### 1.8.1.3    cfDNA/ctDNA as a biomarker of cancer

The tumour-derived fraction of circulating cell-free DNA (cfDNA), termed circulating tumour DNA (ctDNA), may comprise between 0.01% and 90% of total cfDNA, and is exquisitely specific, since it can carry genetic alterations concordant with patient's tumour(s)  (Diehl *et al.*, 2008). These alterations may include SNVs, indels, rearrangements, SCNAs, LOH and microsatellite instability (Diaz and Bardelli, 2014). As such, ctDNA can be detected and discriminated from healthy cfDNA using techniques such as quantitative/digital PCR and next-generation sequencing. Circulating tumour DNA is a dynamic biomarker: total ctDNA levels have been shown to correlate with tumour burden across a range of cancer types and then resolve to undetectable levels following complete tumour regression with pharmacotherapy or curative surgery (Diehl *et al.*, 2008; Shinozaki *et al.*, 2007; Forshew *et al.*, 2012; Dawson *et al.*, 2013; Bidard *et al.*, 2014; Hamakawa *et al.*, 2015; Tsao *et al.*, 2015). Therefore, ctDNA levels could serve as a surrogate biomarker for tumour burden to monitor therapeutic response and assess disease status (stable or progressive).

### 1.8.1.4    Quantitative analysis of cfDNA in lung cancer

There has been much interest in quantitative measurement of total cfDNA levels in patients with NSCLC.  For example, several studies have shown a significant difference in total cfDNA levels between cancer patients and healthy controls (age-matched smokers and non-smokers) using qPCR or fluorometric assays (Table 1.7) A recent systematic review and meta-analysis calculated a pooled sensitivity and specificity of 81% and 85%, respectively, for the quantification of cfDNA as a diagnostic tool in NSCLC (Jiang *et al.*, 2016). However, many of the studies included in the review did not distinguish between early and late-stage disease, and therefore the sensitivity and specificity of cfDNA quantification in the early disease is unknown. Moreover, two studies in advanced NSCLC have reported no correlation (Nygaard et al., 2014; Lee et al., 2011).

| Author (*et al*) | Year | Cancer subtype | Stage | *n* Cancer | *n* Control | Specimen | Quantification method | Mean cfDNA (ng/mL) Cancer | Mean cfDNA (ng/mL) Control | Significance | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sozzi | 2001 | Mixed | Mixed | 84 | 43 | Plasma | DNA Dipstick™ | 318 | 18 | *P* < 0.05* | Sozzi *et al.*, 2001 |
| Sozzi | 2003 | Mixed | Mixed | 100 | 100 | Plasma | qPCR (*TERT*, 98 bp) | 75.0 | 4.6 | *P* < 0.05* | Sozzi *et al.*, 2003 |
| Xie | 2004 | Mixed | Mixed | 67 | 44 | Plasma | PicoGreen® | 110.7 | 11.6 | *P* < 0.001 | Xie *et al.*, 2004 |
| Herrera | 2005 | Mixed | Mixed | 25 | 62 | Plasma | qPCR (*ACTB*, 99 bp) | 14.6 | 10.6 | *P* = 0.1754 | Herrera *et al.*, 2005 |
| Ludovini | 2008 | Mixed | Mixed | 76 | 66 | Plasma | qPCR (*TERT*, 98 bp) | 60 | 5 | *P* < 0.0001 | Ludovini *et al.*, 2008 |
| Ulivi | 2008 | Mixed | Mixed | 128 | 103 | Serum | qPCR (*GAPDH*, 125 bp) | 48.0 | 8.8 | *P* < 0.001 | Ulivi *et al.*, 2008 |
| Paci | 2009 | Mixed | Mixed | 151 | 79 | Plasma | qPCR (*TERT*, 98 bp) | 12.8 | 2.9 | *P* < 0.001 | Paci *et al.*, 2009 |
| Szpechcinski | 2009 | Mixed | Mixed | 30 | 16 | Plasma | qPCR (*ACTB*, 100 bp) | 12.0 | 2.7 | *P* < 0.001 | Szpechcinski *et al.*, 2009 |
| Yoon | 2009 | Mixed; SCLC | Mixed | 102 | 105 | Plasma | qPCR (*ACTB*, 99 bp) | 22.6** | 10.4** | *P* < 0.0001 | Yoon *et al.*, 2009 |
| van der Drift | 2010 | Mixed | Mixed | 46 | 21 | Plasma | qPCR (*HBB*, 71 bp) | 52** | 29** | *P* < 0.001 | van der Drift *et al.*, 2010 |
| Kumar | 2010 | Mixed | Mixed | 100 | 100 | Plasma | PicoGreen® | 95.1 | 74.0 | *P* < 0.001 | Kumar *et al.*, 2010 |
| Roth | 2011 | Mixed; SCLC | Mixed | 35 | 28 | Serum | PicoGreen® | 134.1 | 75.3 | *P* < 0.0001 | Roth *et al.*, 2011 |
| Catarino | 2012 | Mixed | Mixed | 104 | 205 | Plasma | qPCR (*TERT*, 98 bp) | 270 | 122.7 | *P* < 0.0001 | Catarino *et al.*, 2012 |
| Ulivi | 2013 | Mixed | Mixed | 100 | 100 | Plasma | qPCR (*GAPDH*, 125 bp) | 47.2** | 9.2** | *P* < 0.0001 | Ulivi *et al.*, 2013 |
| Chiappetta | 2013 | Mixed | NS | 30 | 26 | Plasma | qPCR (*ACTB*, 68 bp) | 94.5 | 42.8 | *P* < 0.001 | Chiappetta *et al.*, 2013 |
| Szpechcinski | 2015 | Mixed | Mixed | 50 | 40 | Plasma | qPCR (*ACTB*, 100 bp) | 8.0 | 2.3 | *P* < 0.0001 | Szpechcinski *et al.*, 2015 |
| Szpechcinski | 2016 | Mixed | Mixed | 65 | 16 | Plasma | qPCR (*ACTB*, 100 bp) | 10.8 | 2.2 | *P* < 0.0001 | Szpechcinski *et al.*, 2016 |
| Wei | 2016 | AC | IV | 50 | 25 | Plasma | PicoGreen® | 2.2 | 5.0 | *P* < 0.0001 | Wei *et al.*, 2016 |

*significance not specified. **median value. AC, adenocarcinoma; NS, not specified

**Table 1.7 – Summary of quantitative cfDNA analyses in lung cancer patients and non-cancer controls.**

## 1.8.1.5    Qualitative analysis of cfDNA in lung cancer

Numerous proof-of-concept studies first reported the presence of tumour-specific microsatellite alterations in cfDNA of patients with both SCLC and NSCLC (Chen *et al.*, 1996; Sozzi *et al.*, 1999; Gonzalez *et al.*, 2000; Sozzi *et al.*, 2001; Beau-Faller *et al.*, 2003; Andriani *et al.*, 2004; Khan *et al.*, 2004). Activating mutations in *KRAS* were the first SNVs to be detected in the plasma of patients with lung cancer in relation to prognosis (Chong *et al.*, 2007; Gautschi *et al.*, 2007; Camps *et al.*, 2011; Nygaard *et al.*, 2013).

Many cfDNA studies have demonstrated the ability of PCR approaches and next-generation sequencing to detect actionable genomic alterations and follow response to therapy in patients with advanced NSCLC (Mok *et al.*, 2015; Wei *et al,.* 2016; Belic *et al.*, 2016). Digital PCR is widely used to detect ctDNA, owing to its ability to detect low-frequency genetic events in a complex genetic background (Nie *et al.*, 2015). Using digital PCR, Yung *et al.* (2009) detected the *EGFR* exon 19 deletion and the L858R mutation in cfDNA with 92% sensitivity and 100% specificity. Moreover, mutant levels were found to correlate with treatment response. Other studies using modified PCR techniques, such as peptide nucleic acid-mediated clamping and allele-specific PCR have identified *EGFR* mutations in cfDNA from patients with advanced NSCLC, albeit with lower sensitivity (Kim *et al.*, 2013b; Wang *et al.*, 2014b; Jing *et al.*, 2014; Mok *et al.*, 2015; Piotrowska *et al.*, 2015).

More recently, targeted deep sequencing approaches have been employed to identify genetic alterations in cfDNA, including those that confer therapeutic resistance. These alterations can then be followed longitudinally using digital PCR, with potential to detect relapsing disease prior to radiological progression. In a patient with *EGFR*-mutant NSCLC receiving TKI therapy, Murtaza *et al.* (2013) demonstrated emergence of the resistance-conferring T790M mutation, which correlated with disease relapse. Similarly, Ishii *et al.* (2015) detected the T790M mutation in cfDNA with a sensitivity and specificity of 82% and 86%, respectively, and an overall 83% concordance with tumour tissue in 18 patients. Lastly, Newman *et al.* (2014) were first to publish a new deep sequencing

strategy, termed <u>CA</u>ncer <u>P</u>ersonalised <u>P</u>rofiling by deep <u>Seq</u>uencing (CAPP-Seq), in which patient-specific mutations were identified in tumour tissue at ultrahigh sensitivity (0.01%) and subsequently tracked in cfDNA. In patients with advanced NSCLC undergoing pharmacotherapy and/or surgery, they demonstrated a correlation between ctDNA levels and tumour burden and also tracked the *KIF5B-ALK* fusion in a patient undergoing therapy with crizotinib, demonstrating a decrease in allele frequency over time (Newman *et al.*, 2014).

Whilst quantitative and qualitative characteristics of ctDNA have been well studied in patients with advanced solid cancers, relatively few studies have investigated ctDNA in early stage disease. Patients with early-stage disease are likely to have low ctDNA levels and, in turn, low mutant allele frequencies, thus requiring highly sensitive assays and detection methods. Bettegowda *et al.* (2014) reported that in stage I disease across a range of tumour types (n = 49), 47% of patients had detectable ctDNA, rising to 83% in stage IV disease (n = 136). A genome-wide copy number analysis of cfDNA from patients with early lung adenocarcinoma using whole-exome sequencing identified subtle but detectable copy number alterations, which enabled discrimination from healthy cfDNA (Xia *et al.*, 2015).

## 1.9        GEM models: a tool for ctDNA analysis in early disease?

Mouse models are fundamental to our understanding of cancer biology. The genetic and physiological similarities between humans and mice have been exploited to identify novel cancer genes, to study cancer gene function, and to understand the molecular and cellular mechanisms of tumourigenesis (Cheon and Orsulic, 2011). Mouse models also have translational applications in chemoprevention, drug development and biomarker discovery (Kucherlapati, 2012). Plasma proteomic analyses have successfully identified tumour-derived protein signatures in mouse models of breast, colorectal, lung and pancreatic cancers using mass spectrometry (Pitteri *et al.*, 2008; Hung *et al.*, 2006; Hung *et al.*, 2009; Taguchi *et al.*, 2011; Faca *et al.*, 2008).

To date, few studies have utilised mouse models of cancer for biomarker discovery in early disease (Kelly-Spratt *et al.*, 2008). Clinically relevant mouse models could provide valuable insights into the utility of ctDNA as an early biomarker, for several reasons. First, ctDNA analysis in humans is challenging due to intrinsic genetic and environmental heterogeneity between patients. Experimental mice are generated from inbred strains and can be housed in identical environmental conditions, thus minimising such extraneous variability. Second, some mouse models can recapitulate the initiating stages of tumour development in humans, which are virtually undetectable in asymptomatic patients. In addition, the ability to control disease onset permits the acquisition of blood samples at defined time points, before and during the course of tumour progression. Matched tumour tissue can be harvested to determine the stage of disease at the time of blood collection. Importantly, age- and sex-matched wild-type littermates can be compared to experimental mice in order to confirm disease-specific changes in cfDNA.

### 1.9.1      Previous ctDNA analyses in mouse models

Several xenograft model studies have demonstrated that ctDNA can be detected in the circulation of mice bearing human tumours (Table 1.8). These studies were conducted to gain insights into the relationship between ctDNA levels and tumour burden, and to determine whether the ctDNA profile changes during cytotoxic therapy. ctDNA detection and quantification in xenograft models is relatively straightforward since it is of human origin and can be discriminated from mouse cfDNA. The majority of studies have used qPCR to target either single-copy human genes (*ACTB*, *GAPDH*, *HBB* and *PSAT1*) or human repetitive elements (*LINE* and *ALU*) (Rago *et al.*, 2007; Cheng *et al.*, 2009; Thierry *et al.*, 2010; Mouliere *et al.*, 2011; Mouliere *et al.*, 2013; Czeiger *et al.*, 2011; Gorges *et al.*, 2012).

Using PCR to amplify a region of the human β-actin gene, Kamat *et al.* (2006) found that levels of ctDNA in the plasma of mice xenografted with human ovarian cancer cells correlated with tumour burden ($P < 0.01$) The correlation between ctDNA levels and tumour burden was confirmed in other xenograft studies (Rago *et al.*, 2007; Thierry *et al.*, 2010; Mouliere *et al.*, 2011; Mouliere *et al.*, 2013; Czeiger *et al.*, 2011; Gorges *et al.*, 2012). Mouliere *et al.* (2013) used allele-specific PCR to show that levels of the *Kras*[G12V] mutation in the cfDNA of mice xenografted with human colorectal cancer cells increased with tumour burden. Moreover, using primers to distinguish between the mouse WT *Kras* allele and human *Kras*[G12V] allele, the authors confirmed that only the ctDNA fraction increased with tumour burden.

Consistent with the predominantly apoptotic origin of cfDNA, two studies have demonstrated that tumour-bearing mice treated with cytotoxic agents (docetaxel or ganciclovir) exhibit an immediate, transient spike in ctDNA levels, followed by a rapid decline to baseline or below baseline levels (Rago *et al.*, 2007; Kamat *et al.*, 2006). Interestingly, Rago *et al.* (2007) observed the same pattern in mice following surgical resection, suggesting induction of mass apoptosis. These phenomena have also been observed in patients with cancer (Swystun *et al.*, 2011; To *et al.*, 2003).

Integrity of cfDNA has been well studied in human cancer and is frequently shown to be reduced (Madhavan *et al.*, 2014). The most commonly reported method for evaluating cfDNA integrity involves PCR amplifying a genomic region with two primer pairs, producing two differentially-sized amplicons. The abundance of 'long' amplicon relative to 'short' amplicon generates a DNA integrity index (DII). A DII close to 1 implies high integrity, whereas a DII close to 0 implies low integrity (high fragmentation). Using a 290/133 bp amplicon ratio (targeting *ACTB*), Mouliere *et al.* reported a sharp decrease in DII ($0.44 - 0.1$) with tumour weight in a colorectal cancer xenograft model (Mouliere *et al.*, 2011). They subsequently evaluated DII in non-tumour derived (mouse) cfDNA of both xenografted and non-xenografted mice and observed no significant difference, demonstrating that the observed decrease in DII is specific to the ctDNA component.

Whilst xenograft models are valuable for investigating the efficacy of novel therapeutic agents and for studying the ctDNA profile before, during and after therapeutic intervention, they are unsuitable for analysis of ctDNA in the early diagnostic setting for several reasons (Fig 1.7). First, the tumour microenvironment, which is understood to play an important role in tumour development, is not recapitulated due to the impaired immune component (Basu and Herlyn, 2009). Second, cell line-derived tumours lack histopathological characteristics of autochthonous tumours and this may affect biological behaviour. Patient-derived tumour material may overcome this limitation (Marangoni and Poupon, 2014). Third, xenografted tumour cells are genetically and morphologically transformed prior to implantation, and hence, early, *de novo* tumour development cannot be recapitulated.
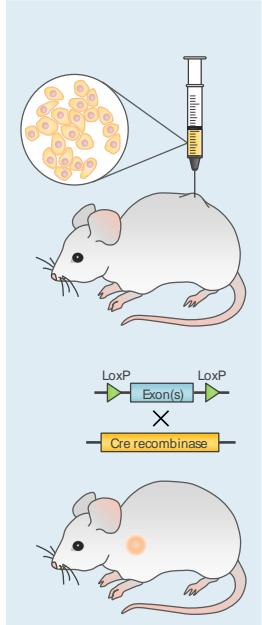
| Author (*et al*) | Year | Xenografted cell line | Tissue of origin | Quantification method | | ctDNA vs tumour burden | Reference |
|---|---|---|---|---|---|---|---|
| | | | | Human (ctDNA) | Mouse (cfDNA) | | |
| Kamat | 2006 | HeyA8 | Ovary | qPCR (*HBB*, 102 bp) | None | $R^2 = 0.8$; $P < 0.01$ | Kamat *et al.*, 2006 |
| Rago | 2007 | CL188<br>HCT116 | Colon<br>Colon | qPCR (*LINE*, 82 bp) | None | Positive correlation (no statistics) | Rago *et al.*, 2007 |
| Cheng | 2009 | A431<br>DLD-1<br>KM12C<br>RPMI-1788<br>SQ5<br>SR-OV-3 | Blood<br>Lung<br>Colon<br>Colon<br>Skin<br>Lung | qPCR (*ACTB*, 100 bp) | qPCR (*ACTB*, 206 bp) | Not studied | Cheng *et al.*, 2009 |
| Thierry | 2010 | HCT116<br>HT29<br>SW620 | Colon<br>Colon<br>Colon | qPCR (*KRAS*, 145 bp; *PSAT1*, 219 bp) | qPCR (*KRAS*, 219 bp; *PSAT1*, 230 bp) | Positive correlation (no statistics) | Thierry *et al.*, 2010 |
| Mouliere | 2011 | HCT116<br>HT29<br>SW620 | Colon<br>Colon<br>Colon | qPCR (*KRAS* 73/145/300 bp; *ACTB*, 133/290 bp) | qPCR (*ACTB*, 214 bp) | Positive correlation (no statistics) | Mouliere *et al.*, 2011 |
| Czeiger | 2011 | MCA-2* | Fibrous | Fluorometric SYBR Gold assay | Fluorometric SYBR Gold assay | Positive correlation, $P = 0.015$ | Czeiger *et al.*, 2011 |
| Gorges | 2012 | COLO 205<br>KPL-4<br>MDA-MB-231<br>MDA-MB-468 | Colon<br>Breast<br>Breast<br>Breast | qPCR (*ALUJ*, 240 bp) | None | Positive correlation (no statistics) | Gorges *et al.*, 2012 |
| Mouliere | 2013 | SW620 | Colon | Allele-specific qPCR (*KRAS*, 189 bp; *BRAF*, 145 bp) | qPCR (*KRAS*, 214 bp) | Positive correlation (no statistics) | Mouliere *et al.*, 2013 |

*mouse cell line

**Table 1.8 - Summary of ctDNA analyses in xenograft models.**

**1.9.2    Potential of GEM models for ctDNA analysis in early cancer**

To date, there have been no publications describing ctDNA analyses in GEM models. These models harbour genetic alterations at endogenous loci and have been pivotal in delineating the roles of specific genes in tumorigenesis. Whereas early models expressed genetic alterations ubiquitously, current models are capable of conditional expression, i.e. in a tissue- and/or time-specific manner, using chemically-induced transcription factors and site-specific recombinases (Frese and Tuveson, 2007). Depending on the cancer type being modelled, oncogenes can be activated, tumour suppressors can be inactivated, and multiple alterations can be induced concomitantly. Some GEM models are effective at recapitulating the early stages of human tumourigenesis, since the tumours are genetically defined, arise *de novo* and interact with a complete tumour microenvironment (Richmond and Su, 2008). Thus, the ability to monitor disease progression at defined stages from initiation permits analysis of ctDNA in early disease. Like xenografts, GEM models also have potential for allowing mechanistic insights into the origins of cfDNA and mechanisms controlling its release. A general disadvantage of the GEM model for recapitulating human tumourigenesis is that tumours are not genetically heterogeneous, since only one or several genetic alterations are introduced. However, one could argue for the purposes of early ctDNA analysis that human early-stage tumours are likely to lack this feature (Fig 1.7)

| | | **Advantages** | **Disadvantages** |
|---|---|---|---|
|  | **Xenograft model** | Technically simple and inexpensive<br><br>Wide choice of source material<br><br>Tumours are genetically complex | Cannot recapitulate early-stage tumour development<br><br>Incomplete tumour microenvironment due to immunodeficiency<br><br>Cell line-derived tumours lack histopathological characteristics of native tumour |
| | **GEM Model** | Can recapitulate early-stage tumour development<br><br>Complete tumour microenvironment<br><br>Genetically well-defined tumours | Genetic complexity of human cancers cannot be recapitulated<br><br>Model development can be expensive and time-consuming |

**Figure 1.7 – Comparison of xenograft models and GEM models for analysis of ctDNA in early disease.**

### 1.9.3      Potential of Braf$^{+/LSL-V600E}$ and Kras$^{+/LSL-G12D}$ models for analysis of cfDNA in early disease

As discussed in Chapter 1.7, the Braf$^{+/LSL-V600E}$ and Kras$^{+/LSL-G12D}$ GEM models are clinically relevant models that exhibit a range of preneoplastic lung lesions that are comparable to those identified in humans.

A key advantage of the Kras$^{+/LSL-G12D}$ mouse model for early ctDNA analysis is the ability to follow tumour progression via micro-computed tomography (µCT) imaging. Kirsch *et al.* (2010) demonstrated a strong correlation between tumour volumes by µCT imaging and histological analysis, which suggests longitudinal imaging may be a non-invasive alternative to harvesting mice at regular time points. This method would significantly reduce animal use and the asynchrony of tumour progression between mice would not complicate analysis since each mouse would serve as its own baseline.

Conditional mutant alleles, such as the *Kras*$^{LSL-G12D}$ allele in the Kras$^{+/LSL-G12D}$ mouse, are latent; the genetic alterations are in the germline, but require Cre-mediated recombination for oncogenic protein expression. Therefore, tumour-derived cfDNA in Cre-Lox-regulated mouse models cannot be defined by the presence of tumour-initiating mutations. Since Cre-Lox recombination leads to oncogene expression and, in turn, malignant cellular transformation, tumour-derived cfDNA is defined by the presence of a single, recombined LoxP sequence (Fig 1.8).
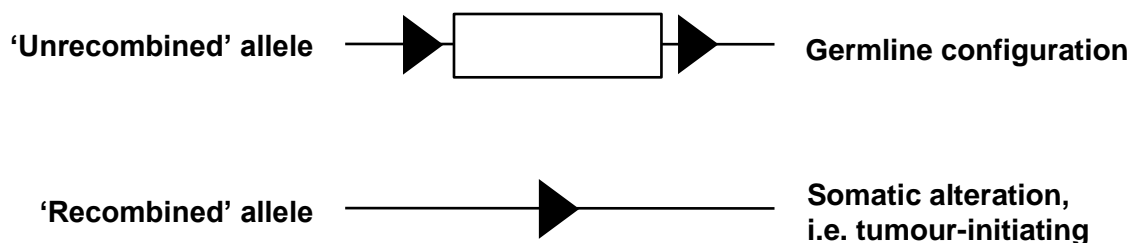


**Figure 1.8 – Tumour and non-tumour origins of LoxP sequences in unrecombined and recombined configurations.** In Cre/Lox-regulated GEM models of cancer, Cre-mediated recombination of LoxP sequences produces a single, 'recombined' LoxP sequence that permits expression of an oncogene or inactivation of a tumour suppressor. Therefore, this single LoxP sequence represents a somatic tumour-initiating alteration that could serve as a biomarker in GEM models of cancer. The white box represents either a transcriptional silencing element in an oncogene or an exon(s) of a tumour suppressor.

## 1.10      Aims and objectives

The work carried out in this thesis comprised two main projects, where:

i)      GEM models of cancer were used to investigate ctDNA in the early, preneoplastic stages of lung adenocarcinoma; and

ii)     SCNAs were evaluated in cfDNA from patients with advanced NSCLC.

The aims of this thesis were to test hypotheses that:

i)      ctDNA is detectable in the blood of mice with preneoplastic disease using the Braf$^{+/LSL-V600E}$ and Kras$^{+/LSL-G12D}$ conditional GEM models of lung adenocarcinoma; and

ii)     SCNAs can be detected in cfDNA from patients with both early and advanced NSCLC.

The objectives were to:

i)      develop a quantitative PCR-based strategy for detection of ctDNA in the Kras$^{+/LSL-G12D}$ mouse;

ii)     determine whether ctDNA can be detected in the Kras$^{+/LSL-G12D}$ mouse at various stages of disease;

iii)    correlate levels of total cfDNA and ctDNA with tumour burden in the Kras$^{+/LSL-G12D}$ mouse;

iv)     analyse SCNA in cfDNA from 183 patients with NSCLC by qPCR; and

v)      compare analysis of SCNA by qPCR and ddPCR.

# Chapter 2

# Materials and methods

## 2.1      Statements of regulatory approval

All animal procedures carried out in this thesis were scrutinised by the University of Leicester Animal Welfare and Ethical Review Body (AWERB) and carried out under UK Home Office License authority (Project Licence holder: Professor Catrin A Pritchard). All laboratory work involving human blood samples was carried out in a Good Clinical Laboratory Practice (GCLP) compliant laboratory.

## 2.2      Patient tissue

All blood samples used in this project were collected in accordance with the Declaration of Helsinki, with written informed consent prior to patient participation. Human plasma samples were obtained from the ReSoLuCENT (Resource for the Study of Lung Cancer Epidemiology in North Trent) study and the GALAXY-1 clinical trial. The ReSoLuCENT study also provided genomic DNA from white blood cells. Samples from the ReSoLuCENT study (NCT00693836), sponsored by the NIHR Clinical Research Network, are referred to as the 'ReSoLuCENT cohort.' The ReSoLuCENT study protocol was approved by West Midlands Multi-centre Research Ethics Committee (05/MRE07/72). Samples from the GALAXY-1 trial (NCT01348126), sponsored by Madrigal Pharmaceuticals Inc. (formerly Synta Pharmaceuticals Corp.) (PA, USA), are referred to as the 'GALAXY-1 cohort.'

## 2.3    Reagents, kits and consumables

All laboratory reagents and consumables are detailed in Tables 2.1 and 2.2. Oligonucleotide primer sequences and assay details are shown in Tables 2.3 – 2.5.

| Item | Supplier | Catalogue # |
|---|---|---|
| Industrial methylated spirit (99% v,v) | Fisher Scientific | 11492874 |
| AccuGENE™ 1X TE buffer | Lonza | 51235 |
| DMEM low glucose without L-glutamine | Lonza | BE12-707F |
| Dulbecco's PBS (without $Ca^{2+}$ and $Mg^{2+}$) | Lonza | 17-512F |
| RPMI (without L-glutamine) | Lonza | BE12-167F |
| Acetic acid | Sigma-Aldrich | A6293 |
| Agarose | Sigma-Aldrich | A9539 |
| Ampicillin (100 mg/mL) | Sigma-Aldrich | A5354 |
| Chloroform:Isoamyl alcohol 24:1 | Sigma-Aldrich | C0549 |
| DPX mountant | Sigma-Aldrich | 44581 |
| DMSO | Sigma-Aldrich | D8418 |
| EDTA | Sigma-Aldrich | 431788 |
| Ethanol (anhydrous, ≥99.5%) | Sigma-Aldrich | 459836 |
| Heat-inactivated fetal bovine serum | Sigma-Aldrich | F9665 |
| HEPES | Sigma-Aldrich | H0887 |
| L-glutamine | Sigma-Aldrich | G7513 |
| LB broth (Lennox) (powder) | Sigma-Aldrich | L3022 |
| NaCl (DNase, RNase, and protease-free) | Sigma-Aldrich | S3014 |
| Paraffin wax (mp, 53 – 57 °C) | Sigma-Aldrich | 327204 |
| Paraformaldehyde (crystalline) | Sigma-Aldrich | P6148 |
| Penicillin-Streptomycin | Sigma-Aldrich | P4333 |
| Phenol:Chloroform:Isoamyl Alcohol (25:24:1) | Sigma-Aldrich | P2069 |
| Proteinase K from *Tritirachium album* | Sigma-Aldrich | P4850 |
| Sodium pyruvate | Sigma-Aldrich | S8636 |
| Tris hydrochloride | Sigma-Aldrich | 857645 |
| Trypan blue | Sigma-Aldrich | 93595 |
| Trypsin-EDTA | Sigma-Aldrich | 59430C |

**Table 2.1 – Suppliers of general laboratory reagents.**

| Item | Supplier | Catalogue # |
|---|---|---|
| Custom MGB/NFQ TaqMan Probes | Applied Biosystems | 4316034 |
| Fast SYBR® Green Master Mix (2X) | Applied Biosystems | 4385612 |
| MicroAmp® Fast Optical 96-Well Reaction Plates | Applied Biosystems | 4346907 |
| MicroAmp® Optical Adhesive Film | Applied Biosystems | 4313663 |
| TaqMan® Fast Universal PCR Master Mix (2X) | Applied Biosystems | 4352042 |
| $K_3$-EDTA Vacutainers | BD Biosciences | 368661 |
| HyperLadder™ 50bp/DNA Loading Buffer (5X) | Bioline | BIO-33054 |
| ddPCR™ Supermix for Probes (No dUTP) | Bio-Rad | 186-3024 |
| ddPCR™ Droplet Reader Oil | Bio-Rad | 186-3004 |
| ddPCR™ Droplet Generation Oil for Probes | Bio-Rad | 186-3005 |
| ddPCR™ 2X Buffer Control Kit | Bio-Rad | 186-3052 |
| DG8™ Cartridges | Bio-Rad | 186-4008 |
| DG8™ Gaskets | Bio-Rad | 186-3009 |
| Pierceable foil heat seal | Bio-Rad | 181-4040 |
| DNA LoBind® Microcentrifuge Tubes, 1.5 mL | Eppendorf | 0030108051 |
| Custom FAM-BHQ-1™ double-dye probe | Eurogentec | PB-DD601-004 |
| LNA®-modified bases | Eurogentec | BA-LN001-004 |
| Qubit® dsDNA HS Assay Kit | Invitrogen | Q32854 |
| Qubit® dsDNA BR Assay Kit | Invitrogen | Q32853 |
| Mouse genomic DNA | Promega | G3091 |
| GeneRead DNA FFPE Kit | Qiagen | 180134 |
| QIAamp® DNA Blood Mini Kit | Qiagen | 51106 |
| QIAamp® DNA FFPE Tissue Kit | Qiagen | 56404 |
| Human genomic DNA | Roche | 11691112001 |
| S-Monovette® (1.6 mL K3 EDTA) | Sarstedt Ltd | 05.1081.001 |
| S-Monovette® needle (22G x 1½") | Sarstedt Ltd | 85.144 |
| Unlabelled oligonucleotides (0.025 µmol, desalted) | Sigma-Aldrich | VC00021 |
| CloneJET™ PCR Cloning Kit | Thermo Scientific | K1231 |
| GeneJET™ Gel Extraction Kit | Thermo Scientific | K0691 |
| GeneJET™ Plasmid Miniprep Kit | Thermo Scientific | K0502 |
| ReddyMix™ PCR Master Mix, 1.5 mM $MgCl_2$ (2X) | Thermo Scientific | AB-0575/DC/LD/A |
| Ad5-CMV-Cre Virus ($8 \times 10^{10}$ PFU/mL) | Viral Vector Core Facility | VVC-U of Iowa-5 |
| Ad5-CMV-ß-gal Virus ($8 \times 10^{10}$ PFU/mL) | Viral Vector Core Facility | None (custom) |

**Table 2.2 – Suppliers of specialist reagents, kits and consumables.**

| Assay target | Oligonucleotide name | Oligonucleotide sequence | Length (nt) | Melting temperature (ºC) | Amplicon length (bp) |
|---|---|---|---|---|---|
| ***Braf*WT** | Forward | GCCCAGGCTCTTTATGAGAA | 20 | 61.7 | 466 |
| | Reverse | AGTCAATCATCCACAGAGACCT | 22 | 62.9 | |
| ***Braf*LSL-V600E** | Forward | GCCCAGGCTCTTTATGAGAA | 20 | 61.7 | 140 |
| | Reverse | GCTTGGCTGGACGTAAACTC | 20 | 62.9 | |
| ***Braf*Lox-V600E** | Forward | GCCCAGGCTCTTTATGAGAA | 20 | 61.7 | 518 |
| | Reverse | AGTCAATCATCCACAGAGACCT | 22 | 62.9 | |
| ***Kras*WT** | Forward | GTCGACAAGCTCATGCGGGTG | 21 | 63.3 | 521 |
| | Reverse | CCTTTACAAGCGCACGCAGACTGTAGA | 27 | 64.5 | |
| ***Kras*LSL-G12D** | Forward | AGCTAGCACCATGGCTTGAGTAAGTCTGCG | 31 | 67.2 | 600 |
| | Reverse | CCTTTACAAGCGCACGCAGACTGTAGA | 27 | 64.5 | |
| ***Kras*Lox-G12D** | Forward | TGACACCAGCTTCGGCTTCCT | 21 | 63.1 | 425 |
| | Reverse | TCCGAATTCAGTGACTACAGATGTACAG | 28 | 58.9 | |

**Table 2.3 – Oligonucleotide sequences for PCR genotyping of Braf+/LSL-V600E and Kras+/LSL-G12D mice.**

| Assay target | Oligonucleotide name | Oligonucleotide sequence | Length (nt) | Melting temperature (ºC) | Amplicon length (bp) |
|---|---|---|---|---|---|
| *Braf*[Lox-G12D] | Forward | TAGGGCTTGCTGTCTTGCTT | 20 | 56.8 | 205 |
| | Reverse | TGCTATACGAAGTTATGTGCAGAAG | 25 | 54.9 | |
| *Kras*[WT/Lox-G12D] | Forward | TCAACAAAGAATACCGCAAGG | 21 | 54.5 | 140 WT |
| | Reverse | CAGCTAATGGCTCTCAAAGGA | 21 | 56.1 | 180 Lox-G12D |
| *Kras*[LSL-G12D] | Forward | CAGTCAACAAAGAATACCGCAAGG | 24 | 57.9 | 117 |
| | Reverse | TCTGTTCCACATACACTTCATTCTCA | 26 | 56.7 | |
| | Probe | TAGCTGTCGACATAACTTCG | 20 | 69.3 | |
| *Kras*[Lox-G12D] | Forward | CCAGTCAACAAAGAATACCGCAAGG | 25 | 60.0 | 93 WT |
| | Reverse | TCTGCATAGTACGCTATACCCTGTG | 25 | 59.2 | 133 Lox-G12D |
| | Probe | TCGACA[T][A][A][C]T[T][C]G[T][A][T]A | 18 | 70.0 | |
| *Gapdh* | Forward | CCTCACAATCTGTCTCACCTTATT | 24 | 55.3 | 113 |
| | Reverse | GACCTCTGTAAGTCCGCTTTG | 21 | 57.2 | |
| | Probe | AGCCTTATTGTCCTCGGGCAT | 21 | 70.4 | |
| *Pseudo-Gapdh* | Forward | GTCGTGGATCTGACGTGCC | 19 | 60.6 | 70 |
| | Reverse | CCTGCTTCACCACCTTCTTGA | 21 | 58.9 | |
| | Probe | ACCTGCCAAGTATGATGA | 18 | 68.0 | |

**Table 2.4 – Mouse-specific oligonucleotide sequences for PCR.** Square brackets indicate positions of locked bases. All probes are 5' 6-FAM™- and 3' MGB/NFQ-labelled, except for *Kras*[Lox-G12D], which is 3' BHQ-1™-labelled.

| Assay Target | Genomic position | Oligonucleotide name | Oligonucleotide sequence | Length (nt) | Melting temperature (ºC) | Amplicon length (bp) |
|---|---|---|---|---|---|---|
| *RPPH1* | chr14:20,811,503-20,811,564 | Forward<br>Reverse<br>Probe | CGGAGGGAAGCTCATCAGTG<br>GACATGGGAGTGGAGTGACA<br>CACGAGCTGAGTGCGT | 20<br>20<br>16 | 63.9<br>63.2<br>71.4 | 62 |
| *MET* | chr7:116,380,060-116,380,134 | Forward<br>Reverse<br>Probe | CCATCCAGTGTCTCCAGAAGTG<br>TTCCCAGTGATAACCAGTGTGTAG<br>TGTGGAGCATACATTAAACCAAA | 22<br>24<br>23 | 63.9<br>63.8<br>71.2 | 75 |
| *EGFR* | chr7:55,209,983-55,210,044 | Forward<br>Reverse<br>Probe | CCAAGGCACGAGTAACAAGCT<br>GGAGGCTGAGAAAATGATCTTCA<br>CGCAGTTGGGCACTT | 21<br>23<br>15 | 64.7<br>62.5<br>69.3 | 63 |
| *PIK3CA* | chr3:178,952,038-178,952,106 | Forward<br>Reverse<br>Probe | TGAGCAAGAGGCTTTGGAGT<br>TTTGTTGTCCAGCCACCAT<br>CATGAAACAAATGAATGATGCACA | 20<br>19<br>24 | 63.7<br>62.3<br>71.5 | 74 |
| *FGFR1* | chr8:38,287,208-38,287,276 | Forward<br>Reverse<br>Probe | AGACTCCGGCCTCTATGCTT<br>TTGACGGAGAAGTAGGTGGTG<br>CGTAACCAGCAGCCCCTC | 20<br>21<br>18 | 64.7<br>63.3<br>74.2 | 69 |
| *HER2* | chr17:37,872,088-37,872,150 | Forward<br>Reverse<br>Probe | ATAACACCCACCTCTGCTTCGT<br>GGTGCGGGTTCCGAAAG<br>CACACGGTGCCCTGG | 22<br>17<br>15 | 65.6<br>62.6<br>71.9 | 63 |
| *SOX2* | chr3:181,713,078-181,713,146 | Forward<br>Reverse<br>Probe | ATGGGTTCGGTGGTCAAGTCC<br>CCTGGAGTGGGAGGAAGAGGT<br>AGGCCAGCTCCAGC | 21<br>21<br>14 | 64.0<br>62.6<br>70.3 | 69 |
| *GAPDH* | chr12:6,535,416-6,535,510 | Forward<br>Reverse<br>Probe | GGCTAGCTGGCCCGATTT<br>GGACACAAGAGGACCTCCATAAA<br>ATGCTTTTCCTAGATTATTC | 18<br>23<br>20 | 64.3<br>63.7<br>70.0 | 95 |

**Table 2.5 – Human-specific oligonucleotide sequences for PCR**. All probes are 5' 6-FAM™- and 3' MGB/NFQ-labelled.

## 2.4 Cell culture

Immortalised cell lines were obtained from the ATCC/ECACC or from stocks in the JS/CAP labs and stored long-term in liquid nitrogen. Cell ampoules were thawed for 1 minute at room temperature and then for 1-2 minutes in a 37°C water bath before resuspension in pre-warmed complete Dulbecco's modified Eagle's medium (DMEM; 10% heat-inactivated fetal bovine serum (FBS); 2 mM L-glutamine; 1% penicillin-streptomycin). Cells were centrifuged for 5 minutes at 300 x g, and the cell pellet was isolated and resuspended with pre-warmed complete DMEM. Cells were counted with a haemocytometer and seeded in 25 $cm^2$ culture flasks at a suitable density. Cells were maintained by incubation at 37°C in a humidified atmosphere with 5% $CO_2$. When a confluency of 70 - 90% was reached, cells were detached from the culture flask with the addition of 0.05% trypsin-EDTA and incubation for 5-10 minutes. Cells were then subcultured at a split ratio of 1:2. Fresh medium was provided daily by removing spent medium and replacing with fresh, pre-warmed medium. Cell stocks were maintained by freezing low passage surplus cells ($2 - 4 \times 10^6$ cells/ml) in 1 ml freezing medium (complete DMEM; 10% heat-inactivated FBS; 10% DMSO) in an isopropanol container at -80°C overnight, followed by storage in liquid nitrogen.

## 2.5 DNA extraction from cell lines

Cultured cells were harvested with 0.05% trypsin-EDTA and spun at 300 x g for 5 minutes to obtain a pellet. The supernatant was discarded and the pellet was washed in PBS and spun again. The washed pellet was resuspended in 0.05 M Tris (pH 8.0)/0.1% SDS and incubated with proteinase K (0.5 mg/ml) at 37°C for 1 hour. An equal volume of phenol:chloroform:IAA (25:24:1) was added, vortexed, and spun for 3 minutes at 1,000 x g, before transferring the aqueous (upper) layer to a clean microcentrifuge tube. An equal volume of chloroform:IAA (24:1) was then added and the sample was vortexed and spun for 3 minutes at 1,000 x g and the aqueous layer transferred to a clean microcentrifuge tube. The

aqueous layer was incubated at -20°C overnight with three times the volume of ice-cold 100% ethanol and one tenth the volume of 1.0 M NaCl. DNA was pelleted by centrifugation at 1,000 x g for 15 minutes at 4°C, washed in 70% ethanol, and respun. The ethanol was discarded and the pellet was left to air-dry before resuspending in an appropriate volume of sterile ultrapure water or sterile TE buffer (10.0 mM Tris [pH 8.0]; 1.0 mM EDTA). DNA concentration was estimated using the Qubit 2.0 fluorometer (Invitrogen) and dsDNA high sensitivity assay kit with a detection range of 10 pg/µL – 100 ng/µL, and DNA was stored for long-term use at 4°C.

## 2.6 Tissue and blood analysis

### 2.6.1 Processing and sectioning of FFPE tissue

Rinsed lung tissue was fixed in ten times the volume of 4% paraformaldehyde solution overnight at RT. After a brief wash with PBS, tissue was dehydrated with 70% (v/v) ethanol and stored at 4°C for preservation. Tissue was further dehydrated with 70% IMS (2 hours), 90% IMS (2 hours) and 100% IMS (8 hours), cleared with chloroform (6 hours), and infiltrated with paraffin wax (4 hours). Tissue processing was automated using a Shandon Citadel 2000 Tissue processor (Thermo Scientific). Tissue was then impregnated with paraffin wax to form a paraffin block. Blocks were cooled to 4°C and cut into 5 µm sections using a Leica RM2135 Microtome (Leica Biosystems). Tissue sections were floated on warm, deionised water and transferred to clean, subbed glass slides, before incubating at 37°C overnight.

### 2.6.2 Haematoxylin and eosin staining of FFPE tissue

Tissue sections were dewaxed twice in xylene for 3 minutes each, and gradually brought to water by passing twice through 99% IMS for 1 minute each and once through 95% IMS for 1 minute. Slides were then rinsed with water for 3 minutes and stained with haematoxylin for 5 minutes. After washing with water for 5 minutes, slides were stained with 1% eosin for 1 minute and rinsed briefly with

water. Sections were gradually dehydrated by passing through 95% IMS for 15 seconds, 99% IMS twice for 1 minute each and xylene twice for 3 minutes each. Stained sections were mounted with coverslips using DPX, viewed under a Leica DM500B microscope (Leica Biosystems) with objectives ranging from x2.5 to x40, and imaged with a DFC420 C digital camera (Leica Microsystems).

### 2.6.3    DNA extraction from FFPE tissue

Tissue was dewaxed in xylene for 3 minutes and then passed into 99% IMS for 3 minutes to remove all traces of xylene. After air-drying, DNA was extracted using the QIAamp® DNA FFPE Tissue Kit, according to the manufacturer's instructions, except the incubation time with proteinase K, which was extended to 48 hours to increase DNA yield. For whole-lung DNA extraction, a cell scraper was used to collect tissue into a microcentrifuge tube. For microdissection, a 25G needle was used to isolate defined regions of tissue under a light microscope by manual scraping. Eluted DNA was stored at 4°C for immediate use or -20°C for later use.

### 2.6.4    Isolation of plasma and lymphocytes from whole blood

Whole EDTA blood was centrifuged for 10 minutes at 1,000 g at 4ºC to isolate plasma. Plasma was then transferred to a clean vessel, avoiding the buffy coat and packed erythrocytes, and spun again for 10 minutes at 2,000 g at 4ºC, before decanting the supernatant into aliquots and storing at -80°C for later use. The buffy coat (concentrated lymphocytes) was also isolated as a source of germline control DNA and stored at -80°C. All contaminated laboratory material was disposed of in a clinical waste container or in an autoclave bag after soaking overnight in Distel laboratory disinfectant. No information is available regarding plasma isolation methodology for the samples in Cohort 1 (Chapter 5.3.3.2) of this thesis.

### 2.6.5    cfDNA extraction from plasma and lymphocytes

cfDNA and leukocyte DNA was extracted using the QIAamp DNA Blood Mini Kit, with the spin method according to the manufacturer's protocol. This kit is based on silica membrane DNA purification technology, which, in brief, comprises a cell lysis step, a size-selective DNA capture on a silica membrane, followed by several washes to remove residual contaminants, and a purified DNA elution step with a low TE buffer (10.0 mM Tris, 0.5 mM EDTA [pH 8.0]). DNA was eluted in 70 µL of elution buffer. Where appropriate, cfDNA was quantified using the Qubit® 2.0 fluorometer (Invitrogen) and high-sensitivity dsDNA Qubit reagents, with a detection range of 10 pg/µL – 100 ng/µL. Genomic DNA from leukocytes was quantified with the broad-range dsDNA Qubit reagents, with a detection range of 100 pg/µL – 1000 ng/µL. Purified DNA was stored at 4°C for immediate use or -20°C for later use.

### 2.6.6    Lyophilisation and reconstitution of cfDNA

For cfDNA samples with limiting concentrations, samples were lyophilised and reconstituted in sterile water to concentrate the sample. LoBind microcentrifuge tubes containing the cfDNA were briefly centrifuged, de-lidded, covered with Nescofilm, and the Nescofilm was pierced with a pipette tip. Samples were then snap-frozen in liquid nitrogen for 10-15 seconds. Tubes were placed in a pre-cooled VirTis BenchTop Pro Freeze Dryer (SP Scientific) and lyophilised at -50°C in a vacuum of 100 mTorr for 3 hours. cfDNA was reconstituted in sterile water at an appropriate volume and stored at 4°C for immediate use or -20°C for later use.

## 2.7         Mouse-specific techniques

## 2.7.1        Generation and genotyping of mice

The $Kras^{+/LSL-G12D}$ mouse was generated in the Tyler Jacks lab and obtained from the Mouse Models of Human Cancers Consortium (MMHCC) Mouse Repository (Tuveson *et al.*, 2004). The $Kras^{LSL-G12D}$ allele was constructed by targeting exon 1 of $Kras^{WT}$ in ESCs with a Lox-STOP-Lox (LSL) targeting vector, followed by selection for homologous recombination-positive clones. The LSL targeting vector comprised two arms; the left arm contained the LSL cassette and the right arm contained mutated exon 1 of *Kras*. The LSL cassette comprised a splice acceptor site, a polyadenylation sequence and a $Pur^R$ selectable marker. This configuration enabled the insertion of the LSL cassette upstream of mutated exon 1, such that in the absence of Cre recombinase, the STOP sequence abrogates transcription of the gene and the protein is not expressed (Figure 2.1A). The $Kras^{LSL-G12D}$ allele was introduced to WT (C57BL6) mice by injecting blastocysts with positive ESC clones. The resulting chimaeric mice were backcrossed with C57BL6 mice to generate heterozygotes.

The $Braf^{+/LSL-V600E}$ mouse was generated by the Pritchard lab and was available in-house (Mercer *et al.*, 2005). The $Braf^{LSL-V600E}$ allele was constructed by targeting exons 14 and 15 of $Braf^{WT}$ in ESCs with a Lox-STOP-Lox (LSL) targeting vector, followed by selection for homologous recombination-positive clones (Tuveson *et al.*, 2004; Mercer *et al.*, 2005). The LSL targeting vector comprised three arms; the left arm contained exon 14, the middle arm contained the LSL cassette and the right arm contained mutated exon 1 of *Braf*. The LSL cassette comprised three LoxP sequences, separating a minigene (exons 15-18 of $Braf^{WT}$) containing a splice acceptor and a $Neo^R$ selectable marker. The minigene and selectable marker both had polyadenylation sequences. This configuration enabled the insertion of the LSL cassette between exons 14 and 15 of $Braf^{WT}$, such that in the absence of Cre recombinase, the minigene enables expression of $Braf^{WT}$ protein (Figure 2.1B). The $Braf^{LSL-V600E}$ allele was introduced to WT (C57BL6) mice by injecting blastocysts with positive ESC clones. The

resulting chimaeric mice were backcrossed with C57BL6 mice to generate heterozygotes.

Genotyping of the *Kras* and *Braf* loci were achieved by PCR of genomic DNA extracted from earclip tissue. Tissue was digested by heating in GNTK lysis buffer (0.01% gelatin, 0.45% Nonidet P40, 0.45% Tween-20, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM-Tris HCl [pH 8.0]) and 100 µg/mL proteinase K to $65^oC$ for 2 hours. The unpurified lysate was used directly for recombination PCR with recombination-specific and non-specific primers (Table 2.4).
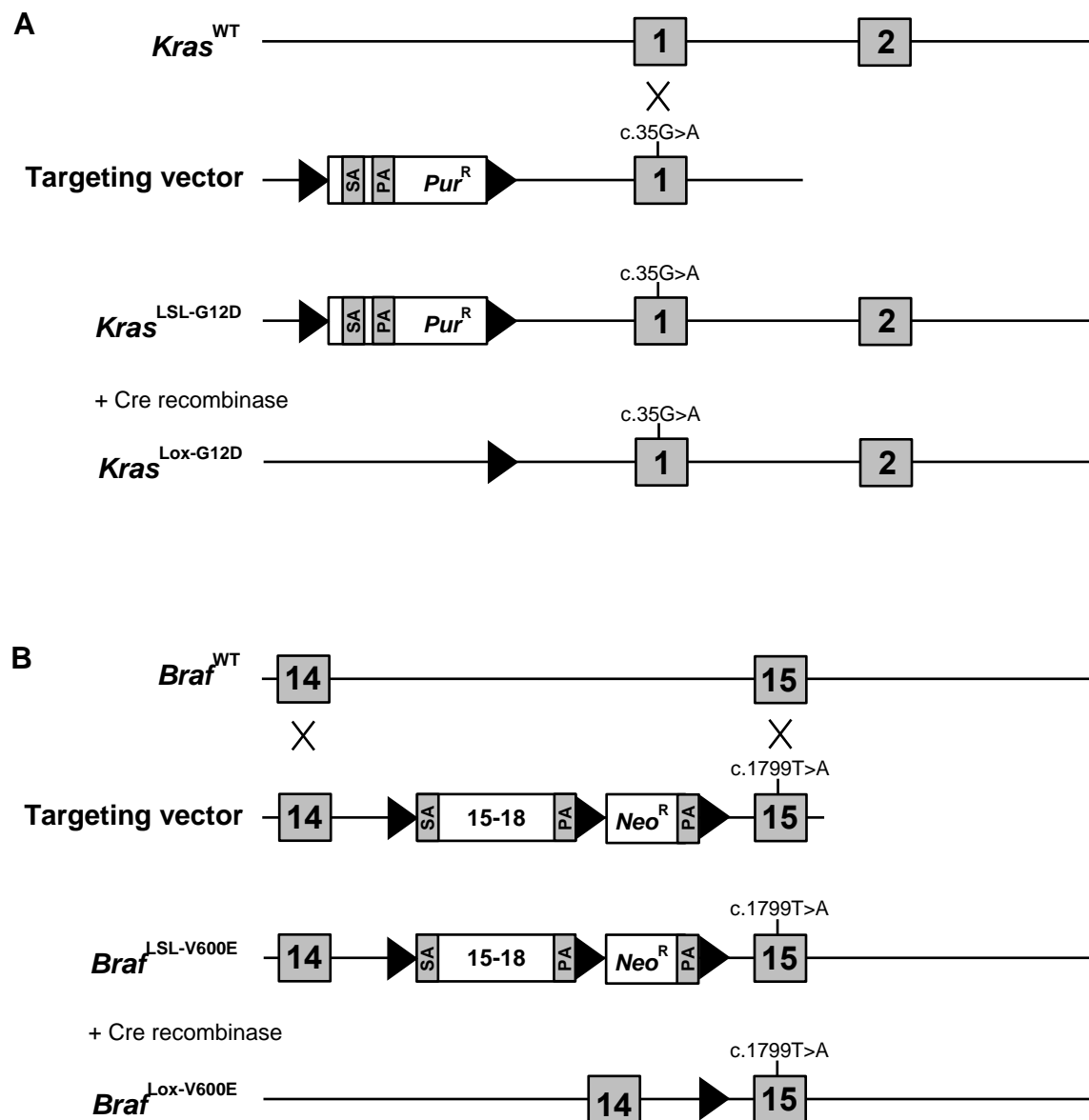
**Figure 2.1 – Strategy for generation of Kras⁺/LSL-G12D (A) and Braf⁺/LSL-V600E (B) mice.**
(A) In ESCs, exon 1 of *Kras*^WT was targeted with a two-arm vector comprising a Lox-STOP-Lox (LSL) cassette and mutated exon 1. In the absence of Cre recombinase, transcription is abrogated at a polyadenylation sequence (PA) within the STOP element. In the presence of Cre recombinase, the STOP element is excised and *Kras*^G12D is transcribed. (B) In ESCs, exons 14 and 15 of *Braf*^WT were targeted with a three-arm vector comprising a minigene (exons 15-18), *Neo*^R gene, exon 14 and mutated exon 15. In the absence of Cre recombinase, exons 15-18 of *Braf*^WT are transcribed from a splice acceptor (SA) in the minigene. In the presence of Cre recombinase, the minigene and *Neo*^R are excised and *Braf*^V600E is transcribed. Black triangles represent LoxP.

## 2.7.2      Expression of conditional *Kras*<sup>Lox-G12D</sup> and *Braf*<sup>Lox-V600E</sup> alleles

The conditional $Kras^{Lox-G12D}$ and $Braf^{Lox-V600E}$ alleles were expressed when mice reached 6 – 8 weeks of age. To achieve lung-specific expression, mice were nasally administered a replication-deficient adenovirus expressing Cre recombinase (adenoviral Cre) under the constitutive CMV promoter (Ad5-CMV-Cre), by staff within the Division of Biomedical Services (University of Leicester). A dose of $5 \times 10^7$ or $1 \times 10^8$ PFU adenoviral Cre (specified in Chapter 4) was suspended in 50 µL sterile PBS and dropped onto the nostrils of the mice under light anaesthesia with 2% isoflurane. Control mice were administered an adenovirus expressing ß-galactosidase (Ad5-CMV-ß-gal) at the same concentration. Viral preparation and administration were carried out in designated class II biosafety cabinets. Adenoviral Cre-treated mice were kept in separate individually ventilated cage systems to ß-gal-treated mice to prevent air-borne transfer of adenovirus between the two groups.

## 2.7.3      Micro-computed tomography (µCT) imaging of the lungs

Mice were imaged under general anaesthesia with an XGI-8 Anaesthesia System (PerkinElmer), consisting of an induction with 3% isoflurane and maintenance with 2 – 3% isoflurane at a flow rate of 1 L/min and in a background of 100% oxygen. Once the breathing rate of the mice dropped below 1 breath/sec and stabilised, µCT data was acquired with a Quantum FX µCT Imaging System (PerkinElmer) using a 40 mm field of view, 80 µm voxel size and an approximate radiation dosage of 58 mGy per 34 sec scan (512 projections in total). Respiratory gating was used to correct for breathing motion-related artefacts. All imaging functions, including automatic 3D reconstruction, were controlled by Quantum FX µCT Software on an embedded computer. Imaging was carried out by Dr Michael Kelly and Ms Justyna Janus (Preclinical Imaging Facility, Centre for Core Biotechnology Services, University of Leicester).

**2.7.3.1    Tumour volume measurement**

Individual tumour volumes were calculated from the reconstructed μCT images using the *Volume Edit* and *Region of Interest* tools in Analyze 11.0 Software (AnalyzeDirect, Inc) according to a pipeline devised by Dr Francois Lassailly (London Research Institute). Before identifying individual tumours, bone structures were identified by defining a voxel intensity threshold (min around + 400; max around + 3000) and extracting the object via the *Extract Object* function. Lung tissue was then identified by defining an empirical voxel intensity threshold. Individual tumours were identified as hyperdense, moderate to well-circumscribed, spherical lesions that appeared approximately round in all slices. Hyperdense lesions that appeared round in some slices and became tubular in others were identified as airway structures. Tumours were defined in each slice using the manual *Draw* tool and the annotated reconstruction was saved as an Object Map. The Object Map file was loaded into the *Region of Interest* tool and the volume of each annotated region (tumour) was calculated. Tumours were first identified in end-point scans and the approximate tissue coordinates were used to track back to previous scans. The μCT analysis pipeline for tumour volume assessment was validated with histological analysis by Mr Callum Rakhit. Kras$^{+/LSL-G12D}$ lungs were μCT imaged and tissue was then FFPE treated, H/E stained and sectioned serially at 20 μm intervals. At each interval, the areas of tumour were calculated using ImageJ Software v1.49 (NIH) and the areas throughout the tumour were stacked to create a 3D estimate of tumour volume.

**2.7.3.2    Functional lung volume measurement**

Functional volumes were calculated from the reconstructed μCT images using the *Volume Edit* and *Region of Interest* tools in Analyze 11.0 Software (AnalyzeDirect, Inc) (Fig 2.2). Image intensity was manually adjusted to optimise contrast between lung airspace and surrounding tissue. Functional airspace was identified as a hypodense mass in the thoracic cavity and segmented with the region-growing method. Using the *Object Extractor* function in the *Volume Edit*

tool, a seed point was placed manually on a region of functional airspace close to the diaphragm, avoiding airway structures. The voxel intensity thresholds were manually adjusted until the green dashed line encompassed the entire lung area, and the object was extracted. To confirm all functional volume was extracted, the entire thoracic cavity was viewed in all anatomical planes (axial, coronal and sagittal), and the process was carried out in triplicate for each scan. The annotated reconstruction was saved as an Object Map. The Object Map file was loaded into the *Region of Interest* tool and functional volume was calculated for both lungs combined.
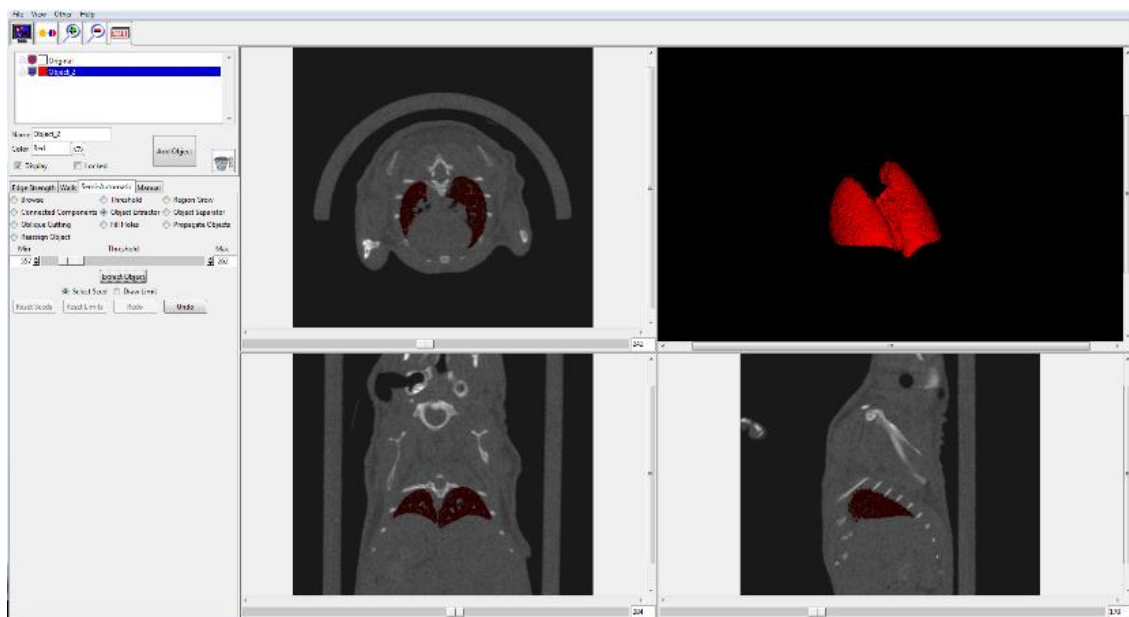


**Figure 2.2 – Measurement of functional lung volume using the *Object Extractor* tool.** Red shading represents areas that have been defined as lung tissue using the region growing method. Top right shows the 3D render of the functional lung volume.

### 2.7.3.3    Total lung volume measurement

Total volumes were calculated from the reconstructed µCT images using the *Volume Edit* and *Region of Interest* tools in Analyze 11.0 Software (AnalyzeDirect, Inc) (Fig 2.3). Image intensity was manually adjusted to optimise contrast between lung airspace and surrounding tissue. The *Auto Trace* tool was used to define the region of interest in each slice. A seed point was placed on a region of lung tissue in each slice and the intensity thresholds were adjusted until the green dashed line encompassed the entire lung area. This process was repeated for all slices in the image stack. The annotated reconstruction was saved as an Object Map. The Object Map file was loaded into the *Region of Interest* tool and total lung volume was calculated for both lungs combined.
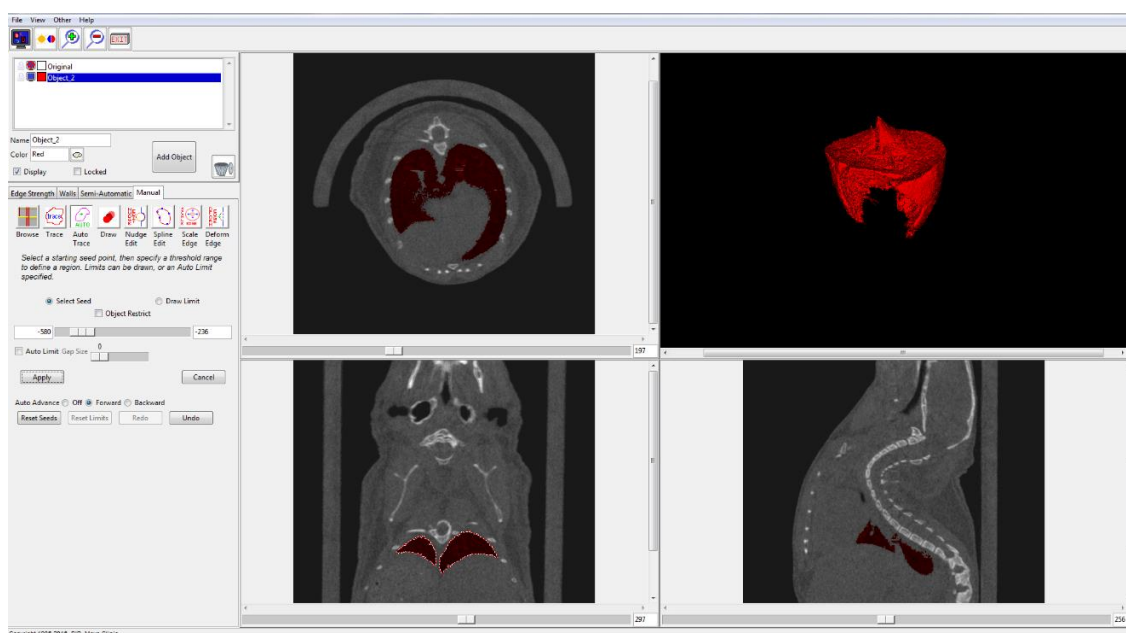


**Figure 2.3 – Measurement of total lung volume using the *Auto Trace* tool.**

### 2.7.4     Collection of mouse blood and tissue

For longitudinal sampling, 40 µL blood samples were taken by saphenous venepuncture at three-week intervals. Blood was aspirated from the leg using an EDTA-coated pipette tip and transferred to 200 µL PBS containing a 4 mM EDTA solution in PBS. Where 40 µL blood was not obtained, the volume of blood was determined by winding down the pipette dial. For terminal bleeds, 200-1000 µL blood was collected by cardiac puncture when either the mice displayed symptoms of the lung cancer phenotype (hyperventilation due to bronchial obstruction and locomotion using back legs due to a net compression of organs towards the back end) or the experiment had reached a pre-defined endpoint. Cardiac blood was collected into 1.6 mL EDTA S-Monovettes using 23G needles and gently agitated to mix the EDTA into the blood.

## 2.8        Analytical techniques

## 2.8.1      Polymerase chain reaction (PCR)

## 2.8.1.1   Assay design, validation and optimisation

Where possible, oligonucleotides were designed using the Primer3 web-based primer design tool (Whitehead Institute, MIT) and aligned to the mouse or human genome using the NCBI nucleotide Basic Local Alignment Search Tool (BLASTN) to confirm specificity. The expected amplicon length for each primer pair was verified with the UCSC *in silico* PCR tool using the mm10 mouse reference assembly or hg38 human reference assembly (Fujita *et al.*, 2011). Sequences were checked for propensity to form hairpins and self/heterodimers using OligoAnalyzer 3.1 (Integrated DNA Technologies), where a ΔG value for secondary structures more positive than - 5 kcal/mol was considered acceptable. Where oligonucleotides could not be successfully designed using Primer3, oligonucleotides were designed empirically by eye. In this case, the following parameters were observed as closely as possible, given the sequence constraints: GC content of 40 - 60%, length of 15 - 30 bp, primer $T_m$ of 58 – 62℃, probe $T_m$ of 68 – 72℃, no runs of > 3 consecutive C or G bases, a GC-clamp containing ≤ 3 G or C bases, and amplicon length < 150 bp. Desalted primers were supplied as lyophilised pellets and reconstituted in sterile $dH_2O$ to a stock concentration of 200 μM. Primers stocks were further diluted to a working concentration of 10 μM.

For probe-based assays, primers were checked for efficiency, specificity and possible primer-dimer formation by SYBR Green melt curve analysis before introducing the probe. The efficiency of each assay was determined by constructing a standard curve from serially diluted genomic DNA (two-fold dilution; seven points starting at 20 ng), with each dilution run in triplicate. Each reaction (10 μL) was performed in triplicate and contained the following: 5 μL Fast SYBR Green Master Mix (2X), 0.6 μL forward primer (600 nM final concentration), 0.6 μL reverse primer (600 nM final concentration), 0.2 μL sterile $dH_2O$ and 3.6 μL template DNA. A no template control (NTC) was included, using sterile $dH_2O$

in place of template DNA. Reactions were set up in MicroAmp Fast Optical 96-well reaction plates, sealed with MicroAmp Optical adhesive film, centrifuged, and run on an Applied Biosystems StepOnePlus™ Real-Time PCR System under fast SYBR Green cycling conditions (enzyme activation at 95°C for 20 seconds, followed by 40 cycles of denaturation at 95°C for 3 seconds and annealing/extension at 60°C for 30 seconds, and enzyme inactivation at 95°C for 15 seconds) and a melt curve (60°C to 95°C with fluorescence readings every 0.3°C).

A standard curve was plotted as log [hgDNA] vs. cycle threshold ($C_t$), and E was calculated as:

$$E = 10^{-\frac{1}{slope}} \qquad\qquad \% E = \left(\left(10^{-\frac{1}{slope}}\right)-1\right) \cdot 100$$

Efficiency between 95% and 105%, and $R^2 > 0.98$ were deemed acceptable. After confirming presence of a single melt peak (i.e. single product and absence of primer-dimers), the qPCR product was run on a 3% agarose gel to confirm amplicon size (Chapter 2.8.3). If the amplicon size determined by agarose gel electrophoresis was as expected, the SYBR Green assay was converted to a TaqMan assay by introduction of a 5' 6-FAM™, 3' MGB/NFQ or 3' VIC®-labelled hydrolysis probe. Assay efficiency was again determined by constructing a standard curve of serially diluted genomic DNA. Since digital PCR does not require near-perfect amplification efficiency, assay optimisation for ddPCR did not involve standard curve generation. To optimise each assay, an annealing temperature ($T_a$) gradient was performed with a range of temperatures above and below the calculated $T_m$ of the primers (typically 55°C - 65°C, over eight temperatures). The optimum $T_a$ for each assay was defined as the temperature which permitted the greatest discrimination between positive and negative droplets whilst avoiding non-specific amplification.

### 2.8.1.2     Manual (end-point) PCR

PCR was performed in a GeneAmp® PCR System 9700 (Applied Biosystems) with ReddyMix PCR Master Mix in a final volume of 25 µL. Each reaction contained 500 nM forward primer, 500 nM reverse primer, 0.625 Units ThermoPrime Taq polymerase and 1.5 mM $MgCl_2$. Generic cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 40 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds. PCR products were resolved by agarose gel electrophoresis (Chapter 2.8.3). Oligonucleotide sequences are detailed in Tables 2.3 and 2.4.

### 2.8.1.3     Quantitative real-time PCR (qPCR)

TaqMan® assays were developed for absolute or relative quantification of targets in cfDNA. Absolute quantification was achieved by constructing a standard curve of serially diluted genomic DNA (1:2 dilution; 10 points) with a range of 5 ng to 40 pg. Each assay comprised unlabelled forward and reverse primers (final concentration, 600 nM) and a 5' 6-FAM™- (or 5' VIC®-), 3' MGB/NFQ-labelled hydrolysis probe (final concentration, 200 nM). Each 10 µL reaction contained 5 µL TaqMan® Fast Universal Master Mix (2X), 1.4 µL primer/probe mix and 3.6 µL template DNA. Reactions were performed in duplicate and a no template control (NTC) was included on each plate with sterile $dH_2O$. Reactions were set up in MicroAmp Fast Optical 96-well reaction plates, sealed with MicroAmp Optical adhesive film, centrifuged, and run on an Applied Biosystems StepOnePlus™ Real-Time PCR System under fast cycling conditions (enzyme activation at 95°C for 20 seconds, followed by 40 cycles of denaturation at 95°C for 1 sec and annealing/extension at 60°C for 20 sec). Oligonucleotide sequences are detailed in Tables 2.3 and 2.4. For relative quantification of targets in cfDNA, the $\Delta\Delta C_T$ method was used as follows:

$$\Delta\Delta C_T = \text{cfDNA} [C_T (\text{target}) - C_T (\text{reference})] - \text{gDNA} [C_T (\text{target}) - C_T (\text{reference})]$$

## 2.8.1.4    Droplet digital PCR (ddPCR)

Each 22 µL reaction contained 11 µL ddPCR Supermix for Probes (No dUTP) (2X), 1.1 µL primer/probe mix (20X), 3 µL template DNA and 6.9 µL sterile $dH_2O$. Reactions were performed in duplicate. Positive and negative controls were included on each plate, along with a no template control (NTC) comprising sterile $dH_2O$. Prior to droplet generation, reactions were set up in MicroAmp Fast Optical 96-well reaction plates in strips of eight wells, lightly sealed with MicroAmp Optical adhesive film, centrifuged, transferred to DG8 ddPCR Cartridges and the cartridges were sealed with DG8 ddPCR Gaskets. Where strips of eight wells were not all filled with reagents, empty wells were filled with 11 µL ddPCR 2X Buffer Control and 11 µL sterile $dH_2O$. 20 µL of each reaction was mixed with 70 µL ddPCR Droplet Generation Oil for Probes using a QX200™ Droplet Generator (Bio-Rad) and the droplet suspensions were transferred to Eppendorf twin-tec 96-Well x 250 µL Semi-Skirted PCR Plates. Plates were sealed with Pierceable Foil Heat Seal at 180°C using a PX1™ PCR Plate Sealer (Bio-Rad). Plates were run on a C1000 Touch™ Thermal Cycler (Bio-Rad) under standard cycling conditions (enzyme activation at 95°C for 10 minutes, 40 cycles of denaturation at 94°C for 30 seconds and annealing/extension at 60°C for 30 seconds, and enzyme inactivation at 98°C for 10 minutes). The ramp rate was unaltered at 100%. Droplets were read with a QX200™ Droplet Reader (Bio-Rad) using ddPCR Droplet Reader Oil and data was analysed using QuantaSoft™ Software v1.7 (Bio-Rad) under Rare Event Detection settings for detection of discrete events or Copy Number Variation (CNV) settings for analysis of copy number. Absolute quantification did not require standard curves.

## 2.8.1.5    Multiplex preamplification PCR

Multiplex preamplification of DNA was achieved by first generating a 10X primer mix, containing 0.3 µM of each primer in $dH_2O$. The preamplification reaction was carried out in a total volume of 25 µL, containing 12.5 µL 2X TaqMan PreAmp master mix, 2.5 µL 10X primer mix and 10 µL DNA (1 ng total). The reaction was performed in a GeneAmp® PCR System 9700 (Applied Biosystems) with the manufacturer's recommended protocol (initial denaturation at 95°C for 10 minutes, 40 cycles of denaturation at 95°C for 15 seconds and annealing /extension at 60°C for 4 minutes, and a final denaturation at 99°C for 10 minutes).

## 2.8.2    Molecular cloning

DNA from recombined lung tissue (FFPE) of a $Kras^{+/LSL-G12D}$ mouse was cloned to compare $Kras^{WT}$ and $Kras^{Lox-G12D}$ alleles for the presence of mutations. A 271 bp region containing the recombined LoxP sequence and 231 bp region corresponding to the WT allele were amplified by PCR.

### 2.8.2.1    Insert DNA preparation

A 1.5 mm Ø core was punched from a FFPE tissue and genomic DNA was extracted using the GeneRead DNA FFPE Tissue Kit according to the manufacturer's instructions. The genomic region of interest was amplified with PCR, except only 25 cycles of PCR were carried out to reduce the impact of PCR-induced mutations. PCR products were resolved on a 2% agarose gel (Chapter 2.8.3) and the desired band was quickly excised with a scalpel over a UVP Chromato-Vue UV transilluminator (302 nm) to minimise DNA damage. The gel slice was weighed and DNA was extracted using the GeneJET Gel Extraction Kit according to the manufacturer's instructions. In brief, the gel slice was dissolved, PCR products were diluted in a binding buffer, bound to a silica-based membrane and centrifuged to remove primers. Salts were removed with a wash buffer and pure DNA was eluted in 10 mM Tris-HCl buffer. DNA was quantified using the Qubit® 2.0 fluorometer (Invitrogen) and high-sensitivity dsDNA Qubit reagents.

### 2.8.2.2    Blunting of insert DNA

Cloning was carried out using the CloneJET PCR Cloning Kit according to the manufacturer's instructions. The DNA blunting reaction (18 µL total) contained 10 µL reaction buffer, 6 µL water, 1 µL DNA blunting enzyme and 1 µL purified PCR product (or water) (15 ng; 0.25 pmol ends) and was prepared on ice. The reaction was incubated at 70°C for 5 minutes.

### 2.8.2.3    Ligation of DNA

The ligation reaction was prepared by adding 1 µL pJET1.2 blunt cloning vector (50 ng; 0.05 pmol ends) and 1 µL T4 DNA ligase to the completed blunting reaction, thus achieving an insert:vector ratio of 5:1. The reaction was prepared on ice and carried out at RT for 5 minutes. A negative control with no insert DNA and a positive control with 971 bp PCR product (supplied with kit) were included to rule out false-positive recombinants and to verify the efficiency of blunting and ligation, respectively. A map of pJET1.2 is shown in Figure 2.4.
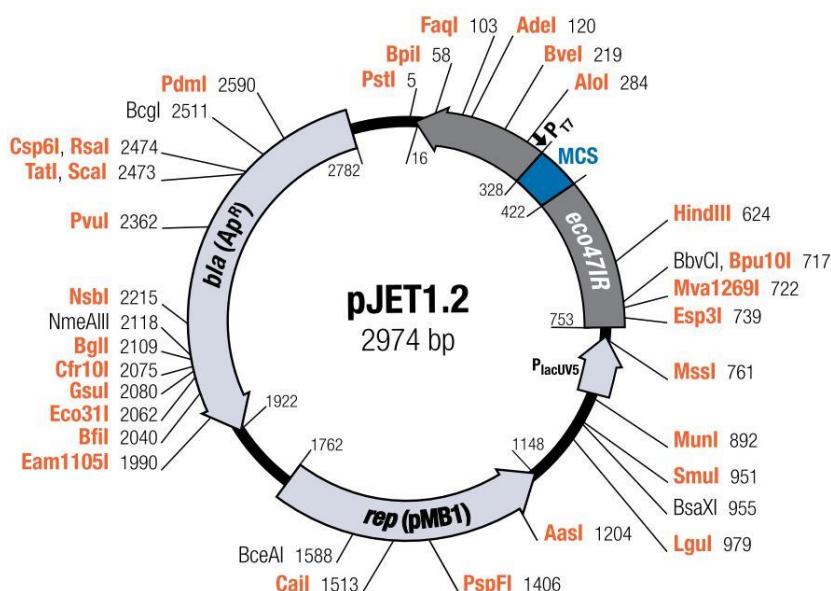


**Figure 2.4 – Map of pJET1.2 blunt cloning vector**. The multiple cloning site (MCS) is located within a lethal gene encoding a restriction enzyme (*eco471R*), which enables positive selection for recombinants. The plasmid also contains a ß-lactamase gene, *bla* (Ap^R), to confer resistance to ampicillin and thus select for transformed cells.

### 2.8.2.4    Transformation of competent *E. coli*

Chemically competent DH5$\alpha$ cells (prepared in-house by Dr Mikael Feracci) were transformed with plasmid DNA via heat-shock. 15 µL competent cells were mixed with 3 µL plasmid DNA and incubated on ice in pre-chilled Eppendorf tubes for 30 minutes, with gentle agitation every 5 minutes. Cells were then heat-shocked in a water bath at 42°C for 30 seconds and returned to ice for 2 minutes. Pre-warmed S.O.C. medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM glucose) was added to a final volume of 250 µL and cells were incubated at 37°C on a shaker (300 rpm) for 60 minutes. Transformed *E. coli* were spread onto two LB agar plates (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 1.2% agar) containing 100 µg/mL ampicillin for colony selection; one plate containing 50 µL cells and the other 150 µL cells. Plates were left to dry, inverted, and incubated at 37°C overnight.

### 2.8.2.5    Liquid culture

Individual colonies of bacteria were picked with a sterile pipette tip and used to inoculate 3 mL aliquots of pre-warmed LB medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) containing 100 µg/mL ampicillin. Liquid cultures were contained in 50 mL falcon tubes and incubated at 37°C on a shaker (200 rpm) for 16 hours. Cells were then pelleted at 5,000 x g for 10 minutes at 4°C.

### 2.8.2.6    Plasmid DNA purification

Plasmid DNA was purified using the GeneJET Plasmid Miniprep Kit according to the manufacturer's instructions. In brief, this began with alkaline lysis of bacterial cells along with RNase A to degrade bacterial RNA, followed by pelleting of cell debris and bacterial DNA. The supernatant was then centrifuged through a silica-based membrane to allow binding of plasmid DNA, the membrane washed to remove salts, and pure plasmid DNA eluted in 10 mM Tris-HCl buffer. Plasmid

DNA was quantified using the Qubit® 2.0 fluorometer (Invitrogen) and broad-range dsDNA Qubit reagents, with a detection range of 100 pg/µL – 1000 ng/µL.

### 2.8.2.7     PCR screening of recombinant plasmids

Prior to sequencing, recombinant plasmids were screened for the presence of the desired insert by PCR (Chapter 2.8.1.2) using the original primers (Table 2.4). PCR products were resolved on a 2% agarose gel (Chapter 2.8.3). Plasmids containing unwanted insert were not taken forward for sequencing.

### 2.8.2.8     Recombinant plasmid sequencing

PCR-verified plasmids containing the desired insert were sent to the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at The University of Leicester for automated dideoxy sequencing in both directions using an Applied Biosystems 3730 Genetic Analyser. The primers were: forward, CGA CTC ACT ATA GGG AGA GCG GC; reverse, AAG AAC ATC GAT TTT CCA TGG CAG. DNA sequence files (.ab1) were analysed with Sequence Scanner v1.0 software (Applied Biosystems).

### 2.8.3     Agarose gel electrophoresis

PCR products were resolved on a 2-3% agarose gel. The gel was prepared by heating an appropriate weight of agarose in TAE buffer (40 mM Tris; 20 mM acetic acid; 1 mM EDTA) until fully dissolved. When cooled to hand-hot (50 – 60ºC), ethidium bromide was added at a concentration of 0.5 µg/mL and the gel was poured into a tray with a comb. Once set, the comb was removed and the gel was placed in an electrophoresis tank containing TAE buffer. DNA was stained with DNA loading buffer (5X) and a 50 bp HyperLadder was included. Gels were electrophoresed at 120 V for approximately 1 hour and imaged with a G:BOX Chemi XX6 Gel Documentation System and GeneSys software (Syngene).

## 2.8.4 Sanger sequencing

Genomic regions of unknown sequence were PCR-amplified and the products were resolved on a 2% agarose gel. The desired product(s) were band-excised with a scalpel and DNA was extracted using the GeneJET Gel Extraction Kit according to the manufacturer's instructions (Chapter 2.8.2.1). Purified DNA was quantified with the Qubit 2.0® fluorometer (Invitrogen) and sent to the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at the University of Leicester for automated dideoxy sequencing in both directions using an Applied Biosystems 3730 Genetic Analyser. DNA sequence files (.ab1) were analysed with Sequence Scanner v1.0 software (Applied Biosystems).

## 2.9 Statistical analysis

Data from qPCR and ddPCR experiments were imported into MS Excel. GraphPad Prism 7.0 (GraphPad Prism Software Inc., CA, USA) was used to generate summary statistics for all quantitative data and to run paired and unpaired t-tests. Cox regression analyses were conducted and Kaplan-Meier plots generated for survival data in MedCalc for Windows (MedCalc Software, Ostend, Belgium). All *P* values in this thesis were considered significant if < 0.05, unless data were subjected to multiple testing. In this scenario (Chapter 5.3.3.1, Figure 5.5), significance levels were adjusted to correct for false discovery rate (FDR) using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

# Chapter 3

# Development of PCR strategies for detection of Cre-Lox recombination in the Kras$^{+/LSL-G12D}$ mouse

## 3.1        Introduction

### 3.1.1        Cre-Lox recombination

The Cre-Lox system of bacteriophage P1 is a site-specific recombination system responsible for circularising the genome and partitioning DNA during reproduction (Shaikh and Sadowski, 1997). The system comprises a 38 kDa *Cre recombinase* protein and two 34 bp *LoxP* recognition sequences. LoxP consists of a core 8 bp sequence flanked by two 13 bp inverted repeat sequences (Fig 3.1A). Cre recombinase catalyses recombination between two LoxP sequences, producing a single, identical LoxP sequence. The location and orientation of one LoxP sequence relative to another determines whether the recombination event leads to a deletion, translocation or inversion of DNA (Fig 3.1B) (Aranda *et al.*, 2001).

Since demonstrating the recombination system can also function in mammalian cells, the Cre-Lox system has become an established tool in the generation and manipulation of genetically engineered mice (Frese and Tuveson, 2007). In cancer models, oncogenes can be activated and tumour suppressor genes can be inactivated by inserting a LoxP-flanked STOP element upstream of the gene, or LoxP-flanking a critical exon, respectively (DuPage *et al.*, 2009). These conditional genetic changes only occur upon delivery of Cre recombinase, which can be controlled spatially using tissue-specific expression and/or temporally using modified forms of the enzyme which become activated by inducing agents (e.g. doxycycline and tamoxifen) (Frese and Tuveson, 2007).

**A**

```
5'  ATAACTTCGTATA GCATACAT TATACGAAGTTAT  3'
3'  TATTGAAGCATAT CGTATGTA ATATGCTTCAATA  5'
```

**B**
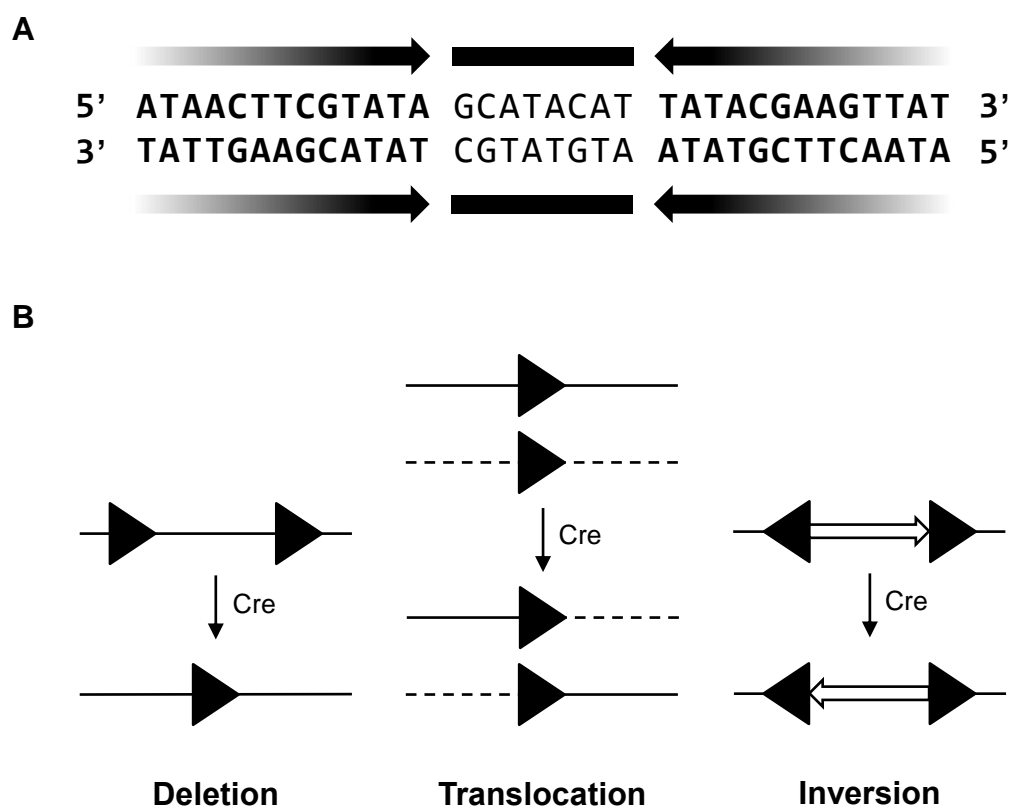


**Figure 3.1 – Cre-Lox recombination**. (A) LoxP is a 34 bp palindromic sequence comprising an 8 bp core region flanked by 13 bp inverted repeats. (B) In the presence of Cre recombinase, LoxP sequences recombine to produce deletions, translocations or inversions of DNA, depending on sequence location and orientation.

## 3.1.2      PCR detection of LoxP

### 3.1.2.1      Genomic DNA

Genotyping of Cre-Lox-regulated mice is achieved by PCR or Southern blot targeting of LoxP sequences in genomic DNA, which is commonly derived from ear or tail tissue. PCR genotyping requires primers that flank a single LoxP sequence or a floxed DNA sequence, such that presence/absence of LoxP can be determined by analysis of product size. This technique can also distinguish between unrecombined and recombined LoxP sequences. Fig 3.2 exemplifies a genotyping strategy for a floxed allele. Weis *et al.* have developed a probe-based qPCR assay to assess for Cre-Lox recombination, which is more sensitive and less time-consuming than end-point PCR (Fig 3.3) (Weis *et al.*, 2010).
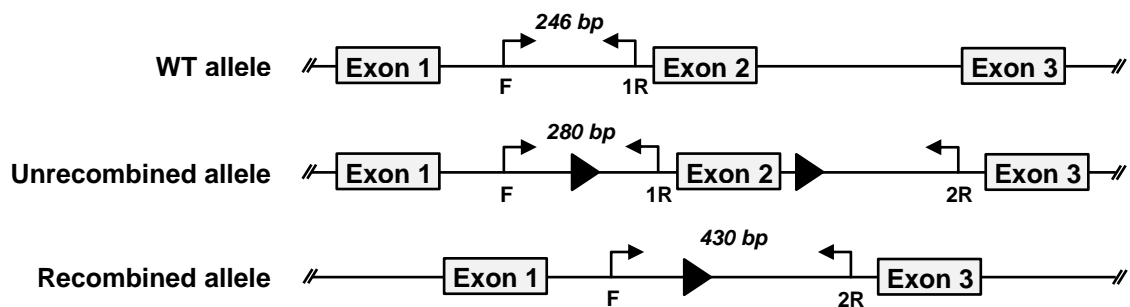


**Figure 3.2 – Example of PCR genotyping for a floxed allele**. PCR primers F and 1R amplify a region of the unrecombined allele and the WT allele, with a product size difference of 34 bp. Primers F and 2R amplify the recombined allele only; the unrecombined allele is too long.
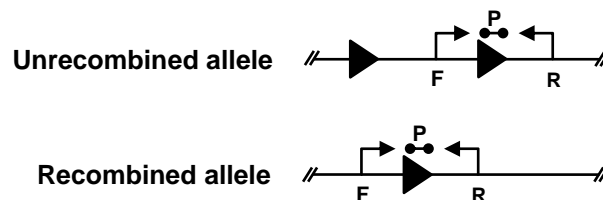


**Figure 3.3 – qPCR assay for assessment of Cre-Lox recombination**. Allele-specific primers flank the LoxP sequence. A universal hydrolysis probe is directed against the LoxP sequence in both alleles. Adapted from Figure 1, Weis *et al.*, 2010.

### 3.1.2.2 Circulating cell-free DNA

There are currently no published studies assessing Cre-Lox recombination in cfDNA. Since many mouse cancer models are driven by conditional alleles under Cre-Lox control, recombination represents a tumour-specific alteration. Thus, quantitative analysis of Cre-Lox recombination in cfDNA from these mice may provide insights into the utility of cfDNA as a biomarker for the early detection of cancer and possibly a biomarker for response to therapeutic agents. There are also no published studies assessing total cfDNA levels over time in GEM models of cancer. Since cfDNA levels are known to increase over time in patients with cancer and in xenograft models, this phenomenon is worthy of investigation in GEM models, and may permit correlation of cfDNA levels with tumour burden.

The design of a PCR strategy for quantitative analysis of Cre-Lox recombination in cfDNA is accompanied by several constraints. First, cfDNA is generally of low integrity and so the PCR amplicon must be kept short. Second, cfDNA exists at low concentrations and the volumes of blood obtainable from a mouse are small, so the assay must be highly sensitive. Based on these considerations, a qPCR approach similar to that of Weis *et al.* may be suitable.

## 3.2        Aims and objectives

The aims of this chapter were to develop PCR approaches for quantification of:

i)        total cfDNA (mouse); and

ii)       Cre-Lox recombination in cfDNA of the Kras$^{+/LSL-G12D}$ mouse.

The objectives of this chapter were to:

i)        design and optimise a qPCR assay for quantification of total mouse cfDNA;

ii)       map the cloning-in position of the LSL element in *Kras*$^{LSL-G12D}$;

iii)      clone and sequence the recombined LoxP site in the Kras$^{+/LSL-G12D}$ mouse to assess for sequence variation; and

iv)       design, validate and optimise dPCR assays for quantification of unrecombined (*Kras*$^{LSL-G12D}$) and recombined (*Kras*$^{Lox-G12D}$) alleles in cfDNA of the Kras$^{+/LSL-G12D}$ mouse.

## 3.3 Results

Work presented in this Chapter was motivated by the outcome of proof-of-concept work presented in Chapter 4.3.1.

### 3.3.1 Design of sensitive assay for quantification of total cfDNA

To analyse quantitative changes in the cfDNA profile of mice longitudinally, a sensitive assay for quantification of total cfDNA from small (approximately 40 µL) blood samples was required. A 70 bp TaqMan® qPCR assay was designed to target 44 *Gapdh* pseudogenes (of which there are 285 in the mouse genome). The assay had a minimum linear dynamic range of 2 – 5,000 pg/reaction, an efficiency of 97.1% and $R^2$ of 0.998, as determined by a standard curve of serially-diluted genomic DNA (Fig 3.4).
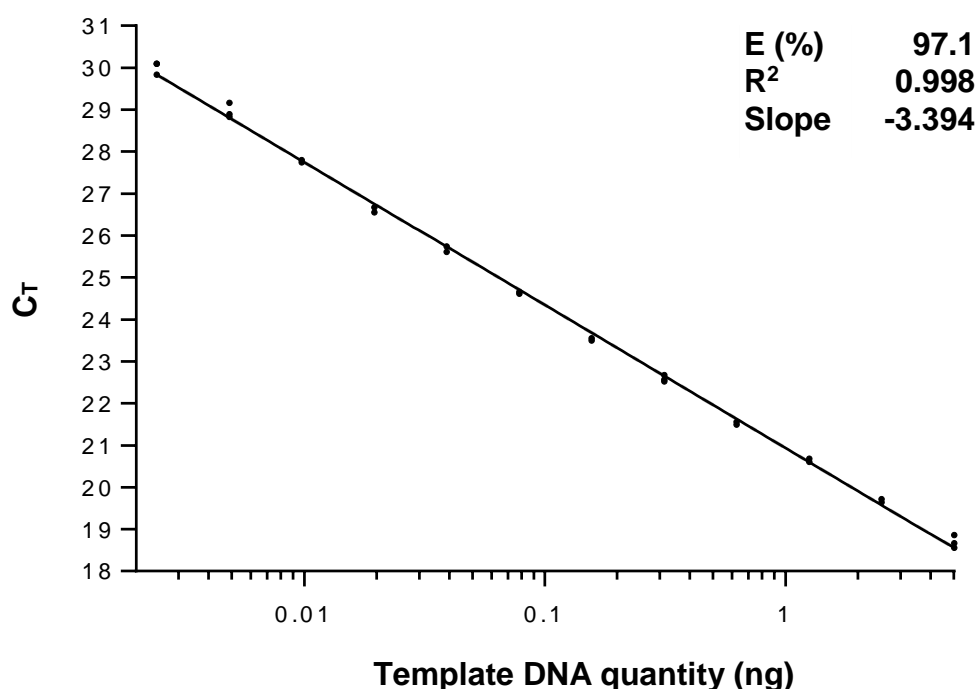


**Figure 3.4 – Standard curve for total quantification of cfDNA by targeting of *Gapdh* pseudogenes.** A TaqMan® qPCR assay targeting 44 *Gapdh* pseudogenes was validated through a 12-point, 1:2 dilution series of genomic DNA. Data points represent three replicates per dilution. $C_T$, cycle threshold.

### 3.3.2    Mapping the position of the LSL element in *Kras*<sup>LSL-G12D</sup>

Detection of *Kras*<sup>Lox-G12D</sup> in cfDNA of Kras<sup>+/LSL-G12D</sup> mice by end-point PCR (Chapter 4.3.1.2) provided justification for characterisation of the recombined LoxP locus within intron 1 of *Kras*<sup>LSL-G12D</sup>, with the view to developing a qPCR assay similar to that of Weis *et al.* (Fig 3.3).

To determine the precise genomic location of the LSL element within exon 1 of *Kras*<sup>Lox-G12D</sup>, the region surrounding the recombined LoxP sequence and the corresponding region of the *Kras*<sup>+</sup> allele were analysed using the 140/180 bp primers in Table 2.4. A schematic of the primer positions on the *Kras*<sup>+</sup> and *Kras*<sup>Lox-G12D</sup> alleles is shown in Figure 3.5. PCR was carried out using recombined MEF cell line DNA as template. PCR products were resolved, gel-purified and subjected to dideoxy sequencing. Pairwise alignment of the two sequences revealed the insert position of the LSL element to be a *Sal*1 restriction site (G^TCGAC) at chr6:145,247,173 – 145,247,178 (GRCm38) (Fig 3.6).
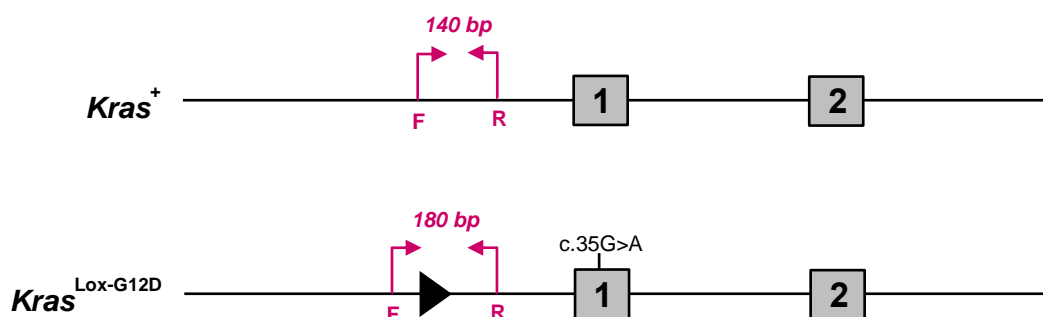


**Figure 3.5 – Positioning of the 140/180 bp primers in *Kras*<sup>+</sup> and *Kras*<sup>Lox-G12D</sup>.** Primers were designed to flank the recombined LoxP sequence in *Kras*<sup>Lox-G12D</sup>, producing a 180 bp amplicon for *Kras*<sup>Lox-G12D</sup> and a 140 bp amplicon for the corresponding region of *Kras*<sup>+</sup>.

**Kras⁺**



**Kras^Lox-G12D**
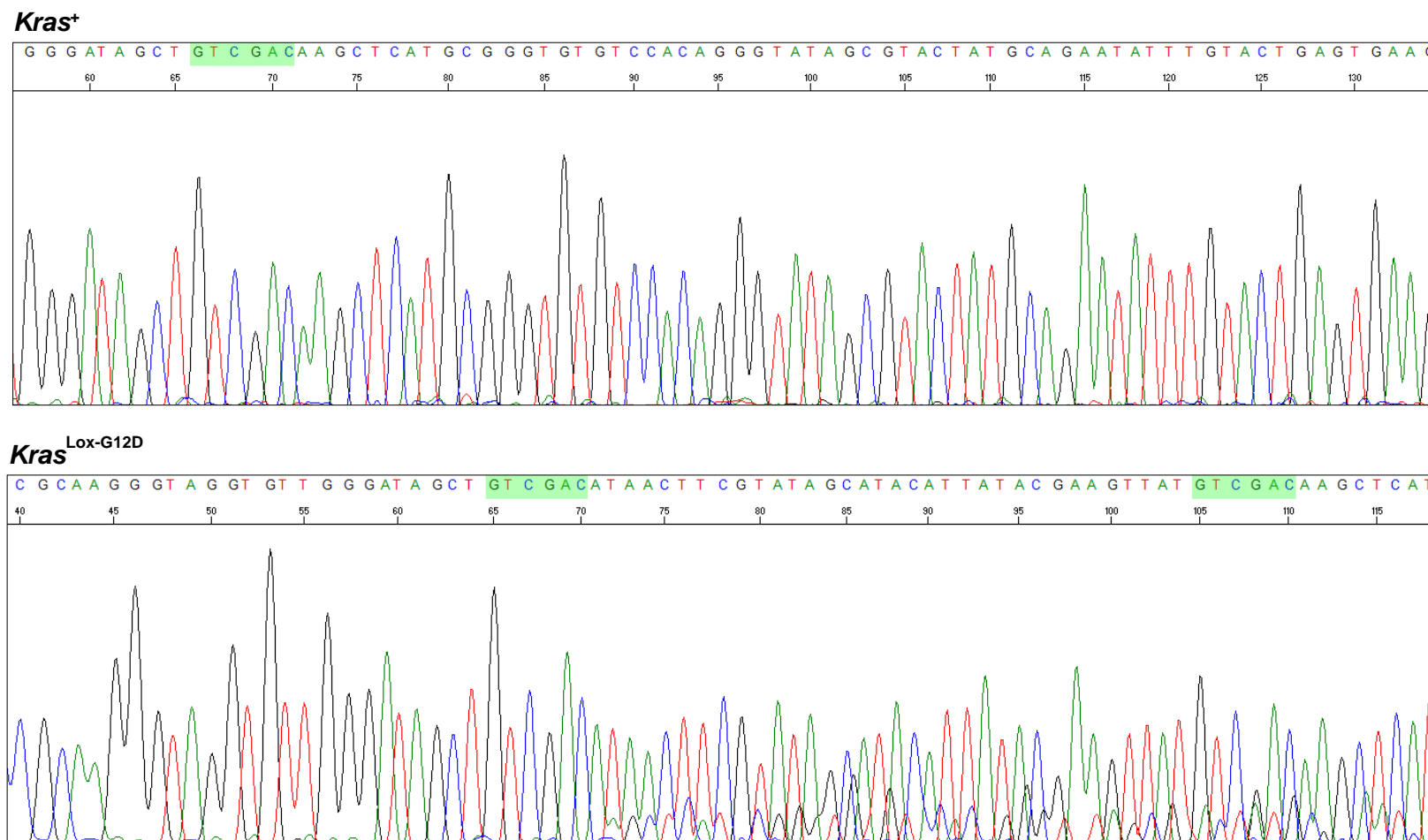


**Figure 3.6 – Electropherogram showing the position of the LSL element in *Kras*^LSL-G12D.** Primers shown in Figure 3.6 were used to amplify the region flanking the recombined LoxP sequence in *Kras*^+/LSL-G12D and the corresponding region of *Kras*⁺. PCR products were sequenced and aligned in a pairwise manner to identify the insert position of the LSL element in *Kras*^LSL-G12D, which was found to be a *Sal*1 restriction site (G^TCGAC) at chr6:145,247,173 – 145,247,178 (GRCm38).

### 3.3.3 Design of assay to detect unrecombined *Kras*

Longitudinal assessment of tumour-derived cfDNA in the Kras$^{+/LSL-G12D}$ mouse requires detection of both unrecombined (*Kras*$^{LSL-G12D}$) and recombined (*Kras*$^{Lox-G12D}$) *Kras* alleles. The unrecombined LoxP sequences fold into predicted stem-loop structures with a length of 13 bp. mFold software (Zuker, 2003) predicted these secondary structures to have a $T_m$ of 55.6°C, which should be sufficiently destabilised to permit PCR amplification at conventional annealing temperatures. Thus, a probe-based qPCR assay similar to that reported by Weis *et al.* was designed (Weis *et al.*, 2010).

The *Kras*$^{LSL-G12D}$ assay was targeted to the left LoxP sequence (most distal to exon 1 of *Kras*$^{LSL-G12D}$). A long minor groove binder (MGB) probe (TAGCTGTCGA CATAACTTCG) was designed to have a high $T_m$ (74°C), ensuring the LoxP stem-loops would have been sufficiently destabilised when the probe annealed. This probe is discussed in Chapter 3.3.5.4 as 'MGB 3' in development of the assay for detection of *Kras*$^{Lox-G12D}$. Specificity for the left LoxP sequence was achieved by placing the forward primer (CCAGTCAACAAAGAATACCGCAAGG) upstream of the LoxP sequence and the reverse primer (TCTGTTCCACATACACTTCATTCT CA) within the STOP region of the LSL element, producing a 117 bp amplicon.

Assay performance was assessed by a generating a standard curve of serially-diluted *Kras*$^{LSL-G12D}$ DNA from a MEF cell line over seven points (Fig 3.7 A). The assay had a minimum linear dynamic range of 160 – 10,000 pg/reaction, an efficiency of 101% and $R^2$ of 0.990. The assay was specific for *Kras*$^{LSL-G12D}$, as shown by the lack of PCR signal for *Kras*$^{+/Lox-G12D}$ and *Kras*$^{+/+}$ templates (Fig 3.7 B, C).

A schematic of the *Kras*$^{LSL-G12D}$ detection assay is shown in Figure 3.8.
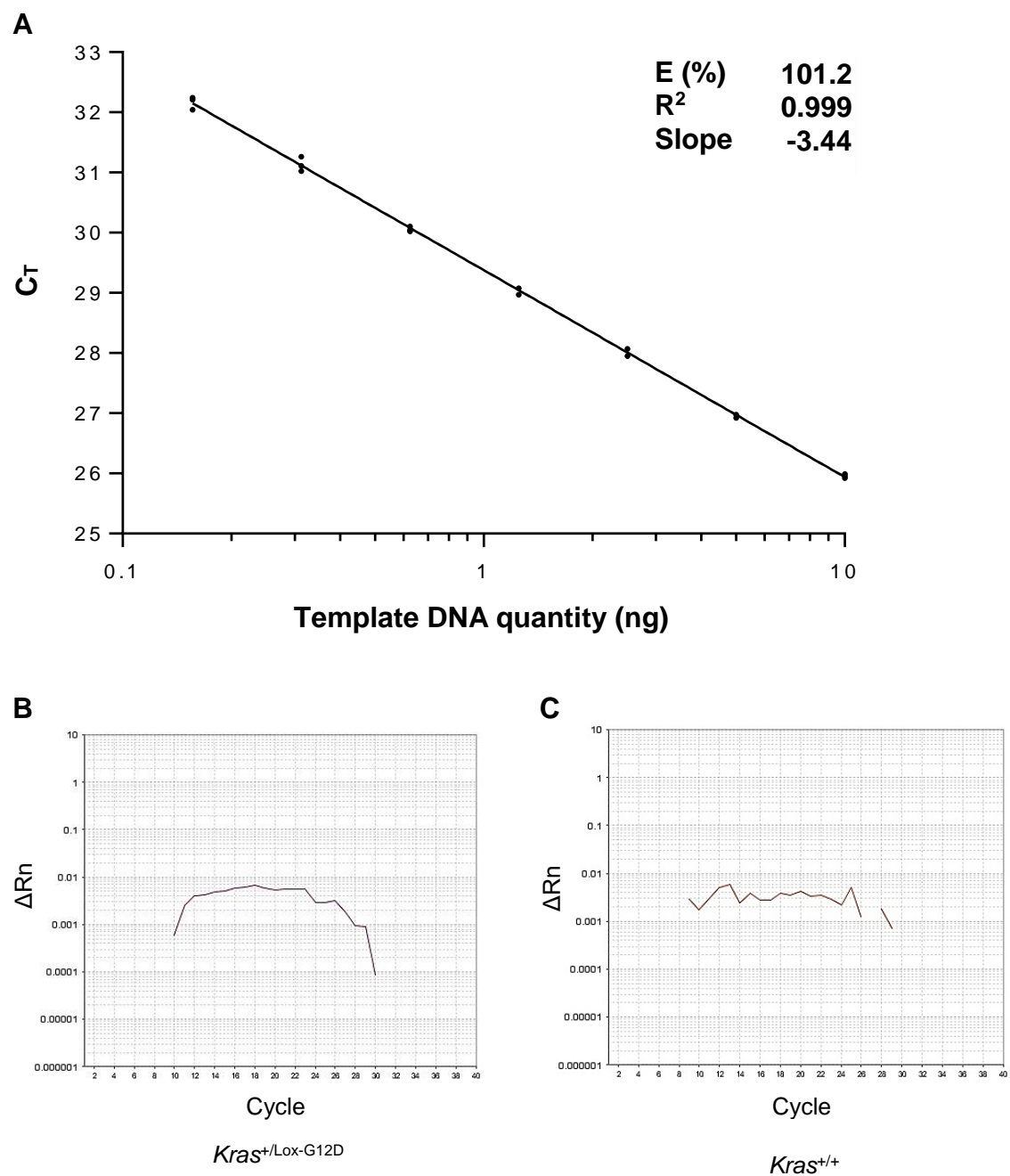
**Figure 3.7 – Performance of qPCR assay for detection of *Kras*<sup>LSL-G12D</sup>.** (A) A 7-point, 1:2 dilution series of *Kras*<sup>+/LSL-DNA</sup> MEF DNA, ranging from 160 pg – 10 ng/reaction. Data points represent three replicates per dilution. $C_T$, cycle threshold. (B) qPCR amplification plot with *Kras*<sup>+/Lox-G12D</sup> MEF DNA. (C) qPCR amplification plot with *Kras*<sup>+/+</sup> MEF DNA.
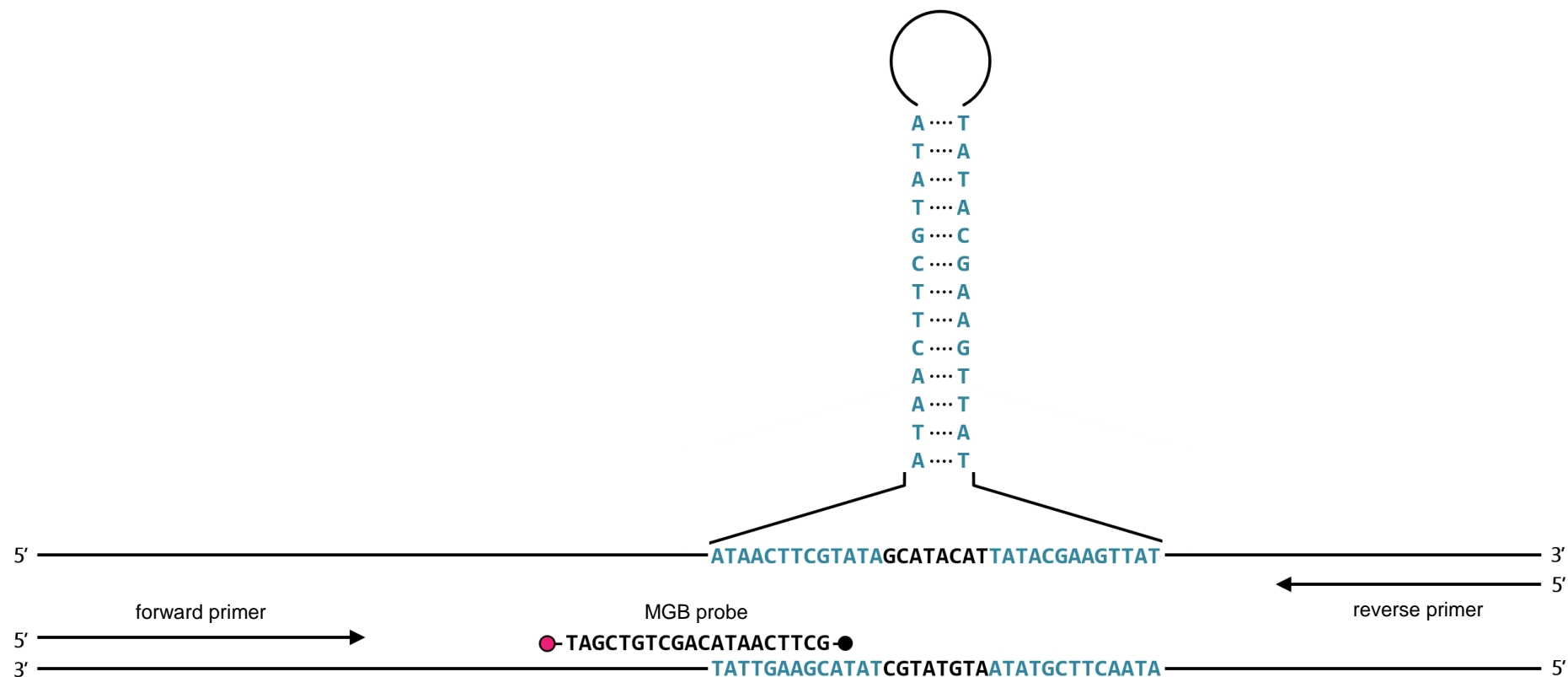
**Figure 3.8 – qPCR strategy for detection of *Kras*<sup>LSL-G12D</sup>.** A 20-base 5' 6-FAM™- 3' MGB-labelled hydrolysis probe targets the right arm of the stem-loop structure on the reverse strand, formed by the LoxP sequence most distal to exon 1 of *Kras*<sup>G12D</sup>. Unlabelled primers flank the unrecombined LoxP sequence, producing an amplicon of 117 bp. The stem-loop on the reverse strand is not shown.

### 3.3.4      Development of an assay to detect recombined *Kras*

#### 3.3.4.1      SYBR Green qPCR

To assess for Cre-Lox recombination in cfDNA of the $Kras^{+/LSL-G12D}$ mouse, several PCR strategies for detection of $Kras^{Lox-G12D}$ (recombined LoxP sequence) were compared. First, a SYBR Green qPCR strategy using the 140 bp/180 bp primer pair was attempted, in which melt-curve analysis could theoretically distinguish between the $Kras^+$ and $Kras^{Lox-G12D}$ amplicons due to the 40 bp difference in length. The predicted melting temperature ($T_m$) values of the 140 bp and 180 bp products were 57.1ºC and 65.7ºC, respectively, as calculated using mFold software (Zuker, 2003). However, generation of a melt-curve from $Kras^{+/+}$ and $Kras^{+/Lox-G12D}$ MEF DNA showed single peaks corresponding to an amplicon $T_m$ of 79.4°C, and resolving the PCR products on an agarose gel confirmed that only the WT (140 bp) amplicon was generated (Fig 3.9).
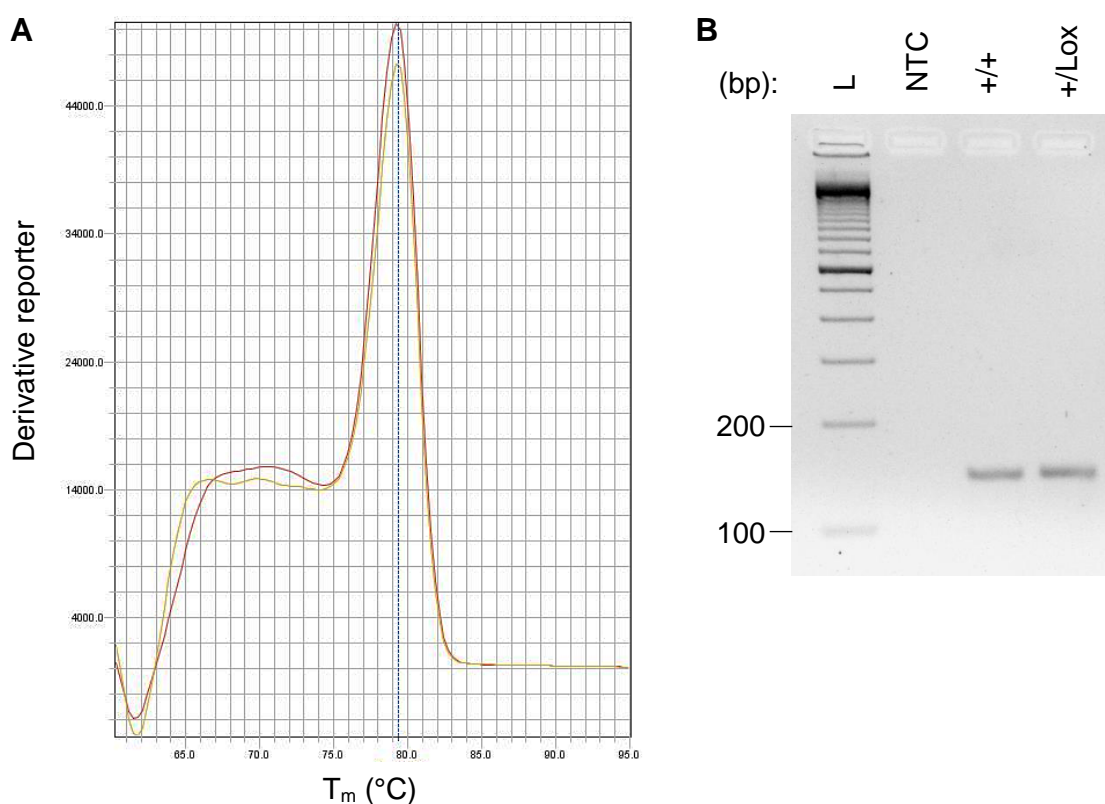


**Figure 3.9 – Melt-curve analysis of PCR products from 140 bp/180 bp primers.** (A) The primers pair that co-amplify $Kras^+$ (140 bp) and $Kras^{Lox-G12D}$ (180 bp) were tested with $Kras^{+/+}$ (red) and $Kras^{+/Lox-G12D}$ (yellow) MEF DNA. (B) Agarose gel from end-point PCR. L, ladder; NTC, no template control.

### 3.3.4.2 End-point PCR strategy with conventional length primers

The result in Fig 3.9 may be explained by the $T_m$ of the primers, which are around 10.4°C lower than that of the *Kras*Lox-G12D PCR product (65.7°C). Another possible explanation for the absence of the *Kras*Lox-G12D (180 bp) product is that amplification of the 140 bp amplicon is favoured due to shorter length.

To prevent co-amplification of *Kras*+, primers were designed to target the recombined LoxP sequence. Multiple primer pairs were designed in which the forward primer and/or the reverse primer extended into the recombined LoxP sequence (Fig 3.10). Primer sequences are given in Table 3.1.
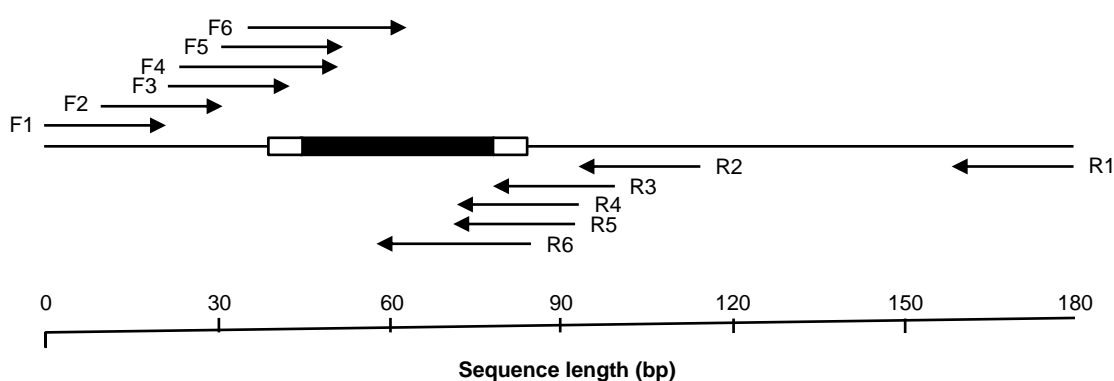


**Figure 3.10 – Targeting of the recombined LoxP sequence in *Kras*Lox-G12D with conventional primers.** To prevent co-amplification of *Kras*+ (as with the 140/180 bp primers, A and G), forward and/or reverse primers were designed to overlap into the LoxP sequence. F, forward; R, reverse.

| Primer | Sequence | Tm (°C) |
|--------|----------|---------|
| F1 | TCAACAAAGAATACCGCAAGG | 59.7 |
| F2 | GAATACCGCAAGGGTAGGTGTT | 61.1 |
| F3 | GTAGGTGTTGGGATAGCTGTCGA | 62.2 |
| F4 | GTGTTGGGATAGCTGTCGACATAA | 62.1 |
| F5 | GGGATAGCTGTCGACATAACTTCG | 62.9 |
| F6 | CTGTCGACATAACTTCGTATAGCATACA | 61.5 |
| R1 | CAGCTAATGGCTCTCAAAGGA | 59.5 |
| R2 | GTGGACACACCCGCATGAG | 63.3 |
| R3 | TGTCGACATAACTTCGTATAATGTATGC | 61.2 |
| R4 | CGCATGAGCTTGTCGACATAAC | 62.6 |
| R5 | GCATGAGCTTGTCGACATAACTTC | 62.0 |
| R6 | CATGAGCTTGTCGACATAACTTCG | 62.8 |

**Table 3.1 – Primer sequences used in the PCR strategy shown in Figure 3.10.**

The strategy shown in Fig 3.10 was carried out by end-point PCR, enabling simple determination of PCR product size using agarose gel analysis. Forward (F1-6) and reverse (R1-6) primers were paired in all 36 possible combinations, and used to amplify from three DNA templates extracted from MEF cell lines: $Kras^{+/+}$ (+/+), $Kras^{+/LSL-G12D}$ (+/LSL) and $Kras^{+/Lox-G12D}$ (+/Lox). PCR products were analysed on an agarose gel. Smeared bands and bands at unexpected sizes were dismissed as primer dimers and non-specific products. Primer combinations that produced visible bands of the expected sizes are shown in Figure 3.11.
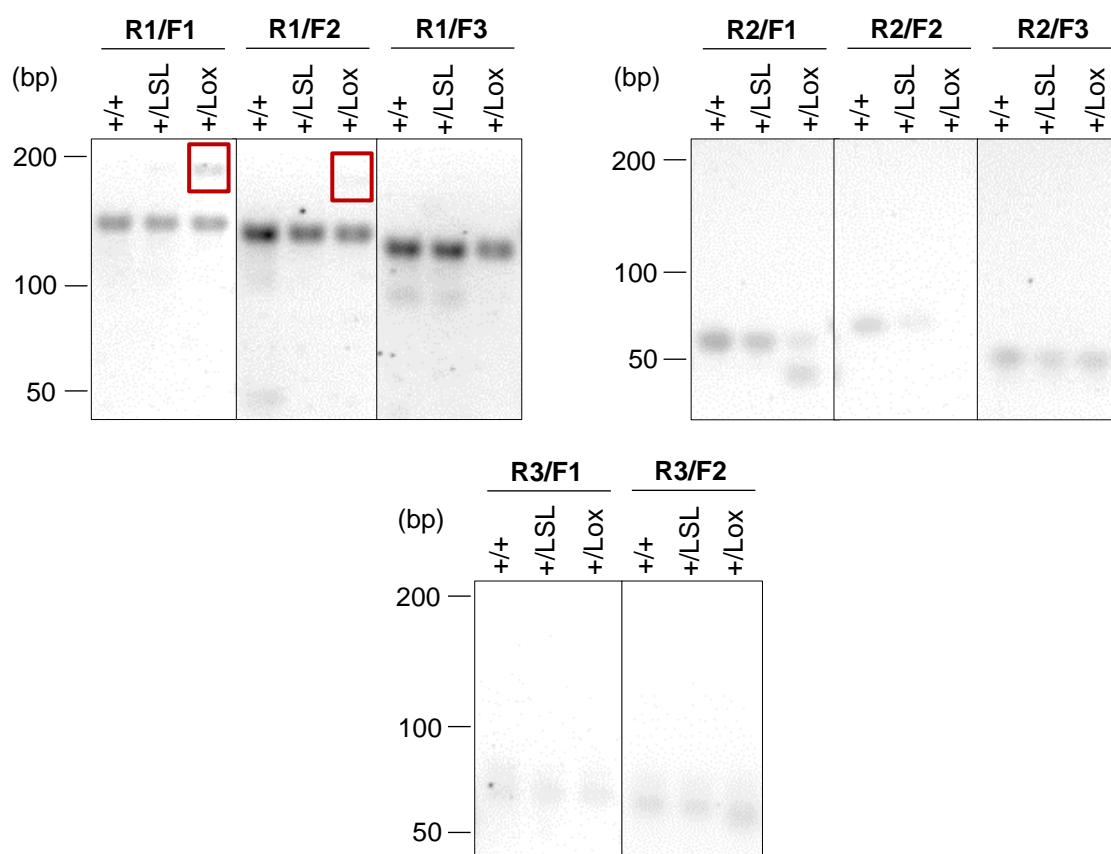


**Figure 3.11 – Agarose gel analysis of PCR products from targeting of the recombined LoxP sequence in *Kras*$^{Lox-G12D}$ with conventional primers.** Primers F1-6 and R1-6) were used in all 36 combinations to amplify *Kras*$^{+}$ and/or *Kras*$^{Lox-G12D}$ from three DNA templates extracted from MEF cell lines: *Kras*$^{+/+}$ (+/+), *Kras*$^{+/LSL-G12D}$ (+/LSL) and *Kras*$^{+/Lox-G12D}$ (+/Lox). Only the primer pairs that generated products of the expected size are shown. Primer positions are shown in Figure 3.17. The bands outlined in red correspond to the *Kras*$^{Lox-G12D}$ product.

As expected, R1 amplified both *Kras*⁺ and *Kras*^Lox-G12D when paired with F1 and F2 (faint bands outlined in red in Fig 3.11). However, only *Kras*⁺ was amplified when R1 was paired with F3. R2 produced amplicons when paired with F1-3 but not F4-6. All amplicons corresponded to *Kras*⁺ and not *Kras*^Lox-G12D. R3 amplified *Kras*⁺, but not *Kras*^Lox-G12D, when paired with F1 and F2. No amplicons were produced when in combination with F3-6.  R4, R5 and R6 did not produce any amplicons when paired with any forward primer. These results show that the WT allele is amplified with primer pairs that do not overlap into the LoxP sequence, but any primers that overlap into this region do not result in amplification of WT (as expected) or *Kras*^Lox-G12D.

### 3.3.4.3    Secondary structure prediction of single-stranded LoxP

PCR strategies so far have demonstrated great difficulty in amplifying across the recombined LoxP sequence using short amplicons suitable for qPCR. Based on the palindromic nature of the 34 bp LoxP sequence, one possible explanation for the failure of previous PCR strategies for amplify across LoxP is that, when single-stranded following PCR denaturation, it folds into a stable stem-loop structure that occludes primer and polymerase access (Huang *et al.*, 2007). In addition, the 6 bp palindromic *Sal*1 restriction sites that flank the recombined LoxP site in the Kras$^{+/LSL-G12D}$ model would result in a 19 bp stem-loop structure with a high $T_m$ (Fig 3.12). Secondary structure analysis of the *Sal*1-flanked recombined LoxP sequence with Mfold Software (Zuker, 2003) predicted the $T_m$ of this structure to be 65.5°C ($\Delta G$ = -12.67 kcal/mol at 37.0°C), assuming a two-state model, which is higher than the $T_a$ of conventional PCR primers in Table 3.1 and may explain failure of the PCR amplicons targeting LoxP.
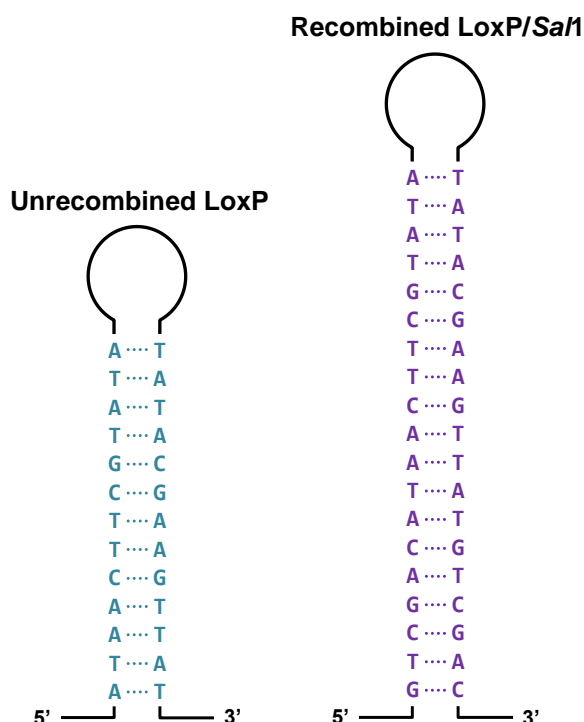


**Figure 3.12 – Stem-loop structure of LoxP in the Kras$^{+/LSL-G12D}$ mouse.** When single-stranded, the 34bp palindromic LoxP sequence folds into a 13 bp stem-loop structure (left) with a predicted $T_m$ of 55.6°C. In the Kras$^{+/LSL-G12D}$ mouse, the LSL element is flanked with 6 bp palindromic *Sal*1 sites, resulting in a *Sal*1-flanked recombined LoxP sequence upon Cre-mediated recombination. When single-stranded, this sequence folds into a 19 bp stem-loop structure (right) with a predicted $T_m$ of 65.5°C. Stem-loops assemble on both strands.

### 3.3.4.4　End-point PCR strategy with high-$T_m$ primers

An alternative strategy was to lengthen the primers in order to increase $T_m$ and, in turn, aid in destabilisation of the predicted 19 bp LoxP/*Sal*1 stem-loop. To this end, long (24 – 41 bp) and high-$T_m$ (66.0 – 77.0°C) primers were designed to flank or overlap into the recombined LoxP sequence, theoretically preventing stem-loop formation at the high annealing/extension temperature of 64.0°C. This strategy is shown in Figure 3.13 and the primer sequences are given in Table 3.2. Figure 3.14 shows the agarose gel analysis of this PCR strategy.
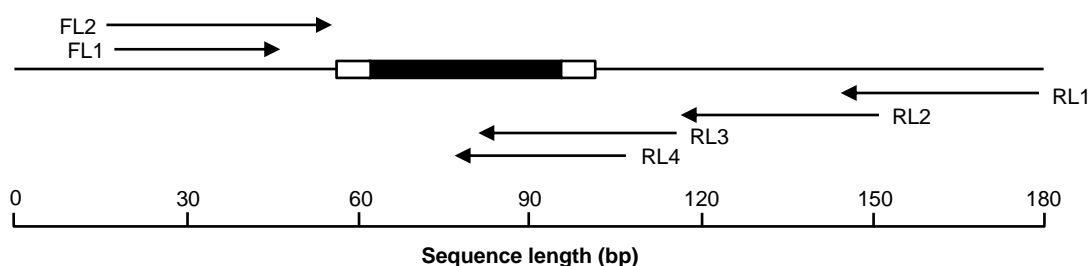


**Figure 3.13 – Targeting of the recombined LoxP sequence in *Kras*^Lox-G12D with high-$T_m$ primers.** Primer positions for forward long (FL) and reverse long (RL) primers targeting the recombined LoxP sequence (black box). Flanking *Sal*1 restriction sites are shown as white boxes.

| Primer | Sequence | Tm (°C) |
|--------|----------|---------|
| FL1 | GAATACCGCAAGGGTAGGTGTTGG | 66.3 |
| FL2 | CAAAGAATACCGCAAGGGTAGGTGTTGGGATAGC | 74.4 |
| RL1 | GGCTCTCAAAGGAATGTATCATGACTTCACTCAGTACAAA | 74.8 |
| RL2 | ACTCAGTACAAATATTCTGCATAGTACGCTATACCCTGTGG | 75.7 |
| RL3 | CACACCCGCATGAGCTTGTCGACATAACTTCGTATAATG | 76.7 |
| RL4 | TGAGCTTGTCGACATAACTTCGTATAATGTATG | 68.8 |

**Table 3.2 – Primer sequences used in the PCR strategy shown in Figure 3.13.**
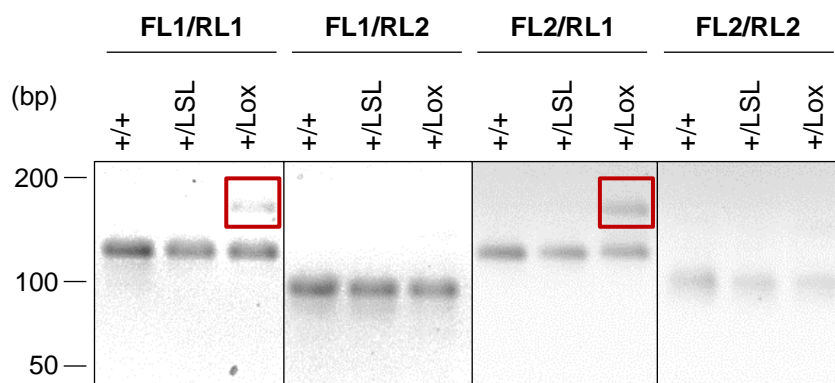
**Figure 3.14 – Agarose gel analysis of PCR products from targeting of the recombined LoxP sequence in *Kras*^Lox-G12D^ with high-T$_m$ primers.** Primers FL1-2 and RL1-4 were used in all 8 combinations to amplify *Kras*$^+$ and/or *Kras*^Lox-G12D^ from three DNA templates extracted from MEF cell lines: *Kras*$^{+/+}$ (+/+), *Kras*$^{+/LSL-G12D}$ (+/LSL) and *Kras*$^{+/Lox-G12D}$ (+/Lox). Only the primer pairs that generated products of the expected size are shown. Primer positions are shown in Figure 3.20. The bands outlined in red correspond to the *Kras*^Lox-G12D^ product.

All primer pair combinations successfully amplified *Kras*$^+$, where neither of the primers overlapped into the LoxP sequence. However, consistent with the previous strategy (conventional length primers), all combinations in which the reverse primer overlapped into the LoxP sequence (RL3, RL4) did not amplify *Kras*^Lox-G12D^. The shortest LoxP-containing amplicon (*Kras*^Lox-G12D^) produced using this high-T$_m$ strategy was 163 bp (123 bp *Kras*$^+$), which was shorter than the amplicon produced with the original 140/180 bp primers, but still too long for quantitative analysis of cfDNA and for use with qPCR.

### 3.3.4.5 TaqMan probe strategy

Another approach was to incorporate a TaqMan probe to improve specific of the PCR. To avoid use of long, non-specific primers, a TaqMan qPCR strategy comprising an MGB hydrolysis probe and conventional primers was developed. Addition of the MGB moiety raised the probe $T_m$ by approximately 10°C, thus allowing a relatively short probe to reach a high $T_m$ without favouring non-specific priming. A high $T_m$ would theoretically permit annealing to (and thus structural destabilisation of) the LoxP sequence before primers anneal.



| | MGB 1 | MGB 2 | MGB 3 | LNA 1 | LNA 2 |
|---|---|---|---|---|---|
| $T_m$ (°C) | 70 | 74 | 72 | 70 | 74 |
| Length | 22 | 20 | 25 | 24 | 18 |
| LNA | 0 | 0 | 0 | 7 | 9 |
| Strand | AS | AS | AS | AS | AS/S |
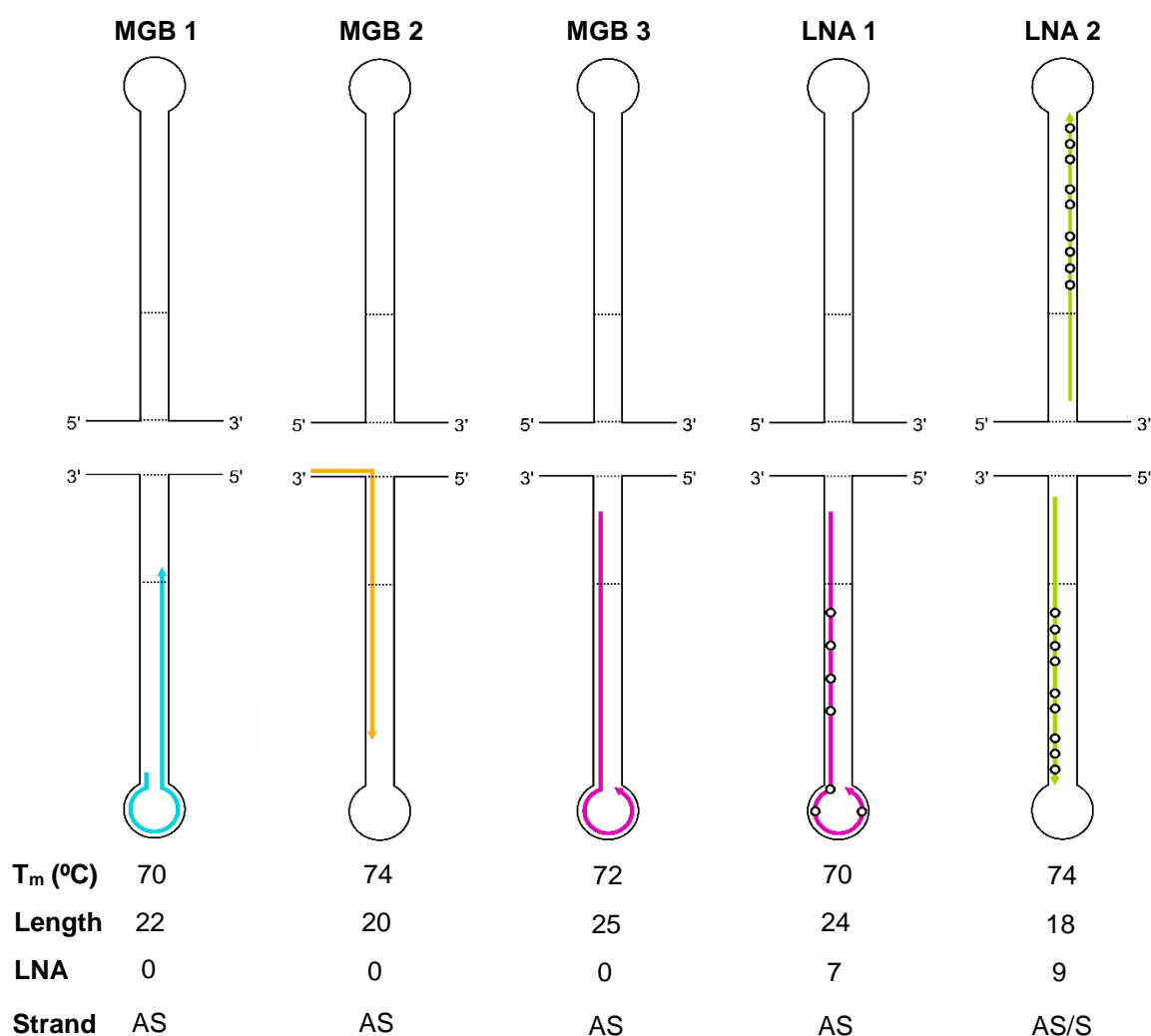
**Figure 3.15 – Targeting of MGB and LNA hydrolysis probes to the recombined LoxP stem loop(s).** Hydrolysis probes targeting the recombined LoxP stem-loop(s) were modified to increase $T_m$; either by addition of a 3' minor groove binder (MGB) moiety or by locking specific bases (represented as hollow circles) using Locked Nucleic Acid (LNA™) technology. Arrows represent 5'→3' directionality of probes.

Three MGB probes ($T_m$ 70.0 – 74.0°C) were designed to destabilise different regions of the stem-loop on the reverse strand (Fig 3.15). MGB 1 was targeted to the 8 bp loop and the left arm of the stem, MGB 2 was targeted to the right arm of the stem and 3' region, and MGB 3 was targeted to the right arm of the stem and 8 bp loop. The probes were not multiplexed. No PCR signal was achieved when each probe was tested with compatible primers from Fig 3.10 (F1 - F5 and R1 - R6 for probe MGB 1; F1 - F2 and R1 - R6 for probes MGB 2 and MGB 3) in all combinations, despite performing a $T_a$ gradient for each primer pair (Fig 3.16).
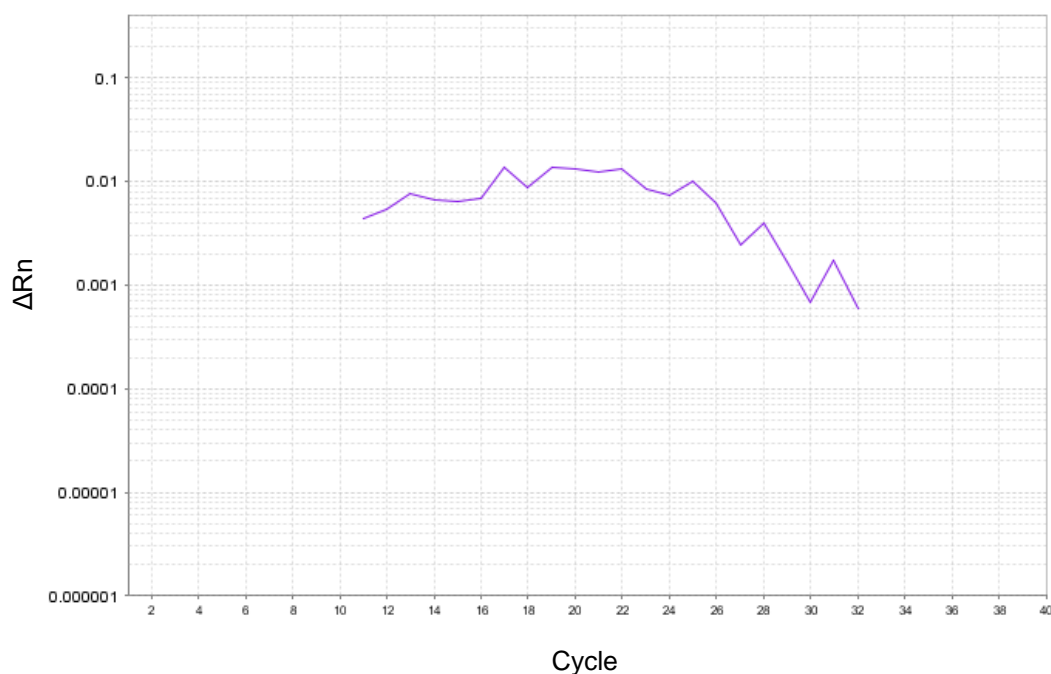


**Figure 3.16 – Representative example of TaqMan qPCR amplification plot for targeting of LoxP sequence in *Kras*[Lox-G12D] with F2/R2/MGB 3 oligonucleotides (no signal).**

### 3.3.4.6    Locked nucleic acid (LNA) probe strategies

After failure of the TaqMan probe strategy to amplify across the recombined LoxP sequence, alternative strategies to provide sufficient destabilisation of the stem-loop were sought. The nucleotide sequence of probe MGB 3 was used to design a locked nucleic acid (LNA) probe (LNA 1, Fig 3.17), in which 7 nucleotides were conformationally restricted by addition of a 2′-O-4′-C-methylene bridge. This 'locked' conformation increases the affinity of oligonucleotides for their complementary sequences, whilst decreasing affinity for highly similar sequences containing minor mismatches. In addition, LNAs are reported to exhibit strand-invasion properties, which could facilitate destabilisation of the LoxP stem-loop (Hertoghs *et al.*, 2003). The modification of probe MGB 3 to incorporate locked bases enabled a $T_m$ of 70.0°C to be reached in the absence of the MGB moiety; at this temperature, the LoxP sequence should not fold into a stem-loop structure. The precise targeting of probe LNA 1 to the LoxP stem-loop is shown in Fig 3.15.

Probe LNA 1 (CGACA[T]A[A]C[T]T[C]GTAT[A]G[C]ATAC[A]T) was first tested in assay format with primers FM 1 (CCAGTCAACAAAGAATACCGCAAGG) and RM 2 (CTGTGGACACACCCGCATGA), which were modified sequences of primers F1 and R2 (Fig 3.10) and had a $T_m$ of 65.8°C and 62.5°C, respectively. In this format, the LNA 1/FM 1/RM 2 assay amplified 113 bp of $Kras^{Lox-G12D}$ and 73 bp of $Kras^+$. A $T_a$ gradient (57.0 – 62.0°C in 1.0°C increments) was carried out with the following DNA templates from MEF cell lines: $Kras^{+/Lox-G12D}$, $Kras^{+/LSL-G12D}$ and $Kras^{+/+}$. The assay demonstrated specificity for the recombined *Kras* allele, since no PCR was obtained from $Kras^{+/LSL-G12D}$ and $Kras^{+/+}$ templates (Fig 3.17).
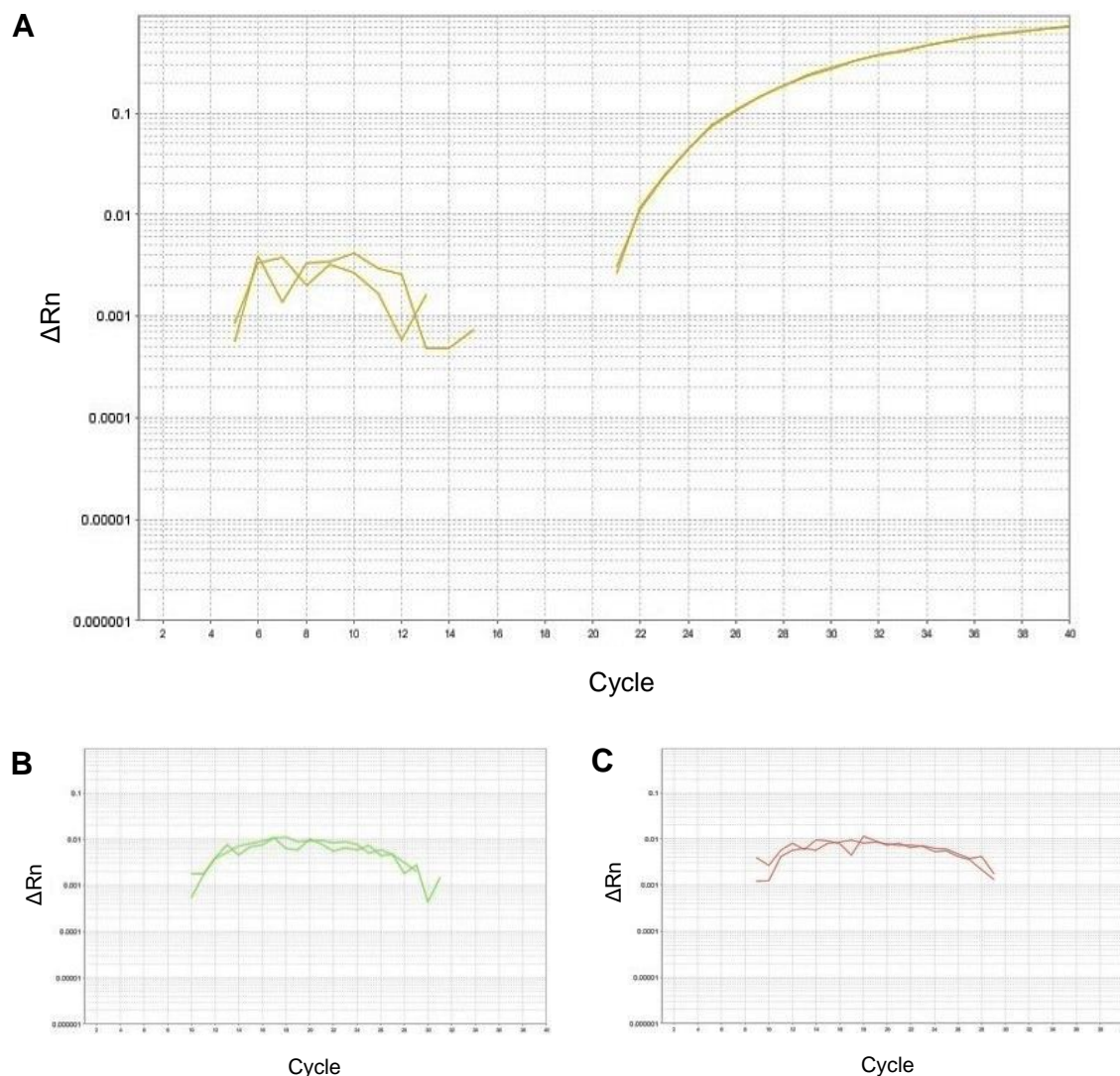
**Figure 3.17 – Targeting of the recombined LoxP sequence in *Kras*<sup>Lox-G12D</sup> with the LNA 1 assay (113 bp).** Probe LNA 1 was tested with primers RM 1 and RM 2 (approximate locations shown in Fig 3.12) and amplification plots at a $T_a$ of 62.0°C are shown. (A) *Kras*<sup>Lox-G12D</sup> MEF DNA. (B) *Kras*<sup>LSL-G12D</sup> MEF DNA. (C) *Kras*<sup>+/+</sup> MEF DNA. Rn, normalised reporter.

To determine the potential utility of the 113 bp LNA 1 qPCR assay (Fig 3.17) in cfDNA, efficiency was calculated by standard curve analysis with serially diluted *Kras*$^{+/Lox-G12D}$ MEF DNA, with a range of 156 pg/reaction – 10 ng/reaction. The assay had 85% efficiency and poor reproducibility (Fig 3.18, solid line).

Since all oligonucleotides in the LNA 1 assay were able to anneal to the same genomic regions of *Kras* in the unrecombined state, efficiency of the assay may be further reduced in the presence of unrecombined allele. To assess the possible influence of *Kras*$^{LSL-G12D}$ on detection of *Kras*$^{Lox-G12D}$ using this assay, a standard curve was generated with *Kras*$^{Lox-G12D}$ MEF DNA as before, but each standard was spiked into a background of *Kras*$^{LSL-G12D}$ MEF DNA (Fig 3.18, dashed line). Frequency of the recombined allele (as % of total unrecombined and recombined) ranged from 50% in the first dilution to 0.78% in the seventh dilution. Whereas the recombined allele could be detected in the seventh dilution in the absence of *Kras*$^{LSL-G12D}$, addition of *Kras*$^{LSL-G12D}$ adversely affected amplification, and the assay failed to detect the recombined allele after the fourth dilution (recombined allele frequency of 6.25%) as shown by the dashed line in Figure 3.18. Therefore, this qPCR assay was not suitable for use in cfDNA.
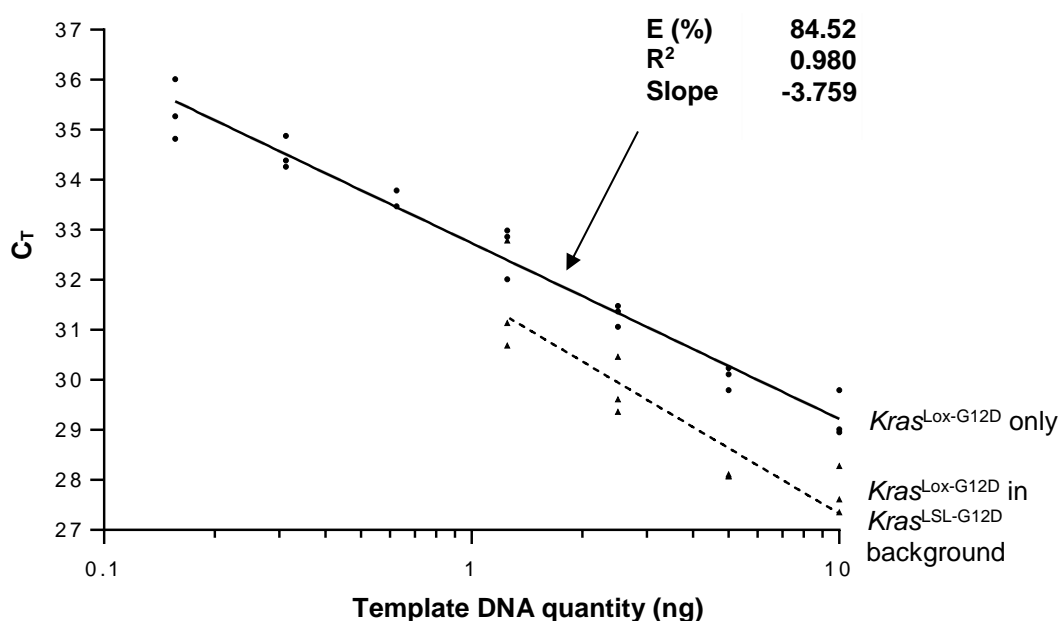


**Figure 3.18 – Standard curve for quantification of *Kras*$^{Lox-G12D}$ in absence and presence of *Kras*$^{LSL-G12D}$ allele with the 113 bp LNA 1 assay.** Efficiency of an LNA probe-based qPCR assay targeting the *Kras*$^{Lox-G12D}$ allele was calculated with two standard curves (7 points, 1:2 dilution). Solid line represents *Kras*$^{Lox-G12D}$ only and dashed line represents presence of *Kras*$^{LSL-G12D}$ allele.

To overcome co-amplification of $Kras^{LSL-G12D}$, the 113 bp LNA 1 qPCR assay was transferred to a droplet digital PCR (ddPCR) system, such that $Kras^{LSL-G12D}$ and $Kras^{Lox-G12D}$ would be unlikely to exist in the same droplets. A $T_a$ gradient (60.0 – 66.0°C over 8 temperatures) was carried out to determine the $T_a$ at which discrimination between positive and negative droplets was greatest, using 10 ng of the following DNA templates from MEF cell lines: $Kras^{+/Lox-G12D}$, $Kras^{+/LSL-G12D}$ and $Kras^{+/+}$. The optimum $T_a$ was found to be 62.0°C (Fig 3.19). In all reactions, negative droplets clustered tightly between 3,000 and 4,000 fluorescence units (FU), but there was additional rain extending up to 5,000; this is not a typical observation with fluorescent probes in ddPCR and was likely a consequence of autofluorescence from the 6-FAM™ fluorophore. The presence of positive droplets in the $Kras^{+/+}$ template reaction suggests occasional mispriming of probe LNA 1 to the *Sal*1 restriction site in $Kras^+$.
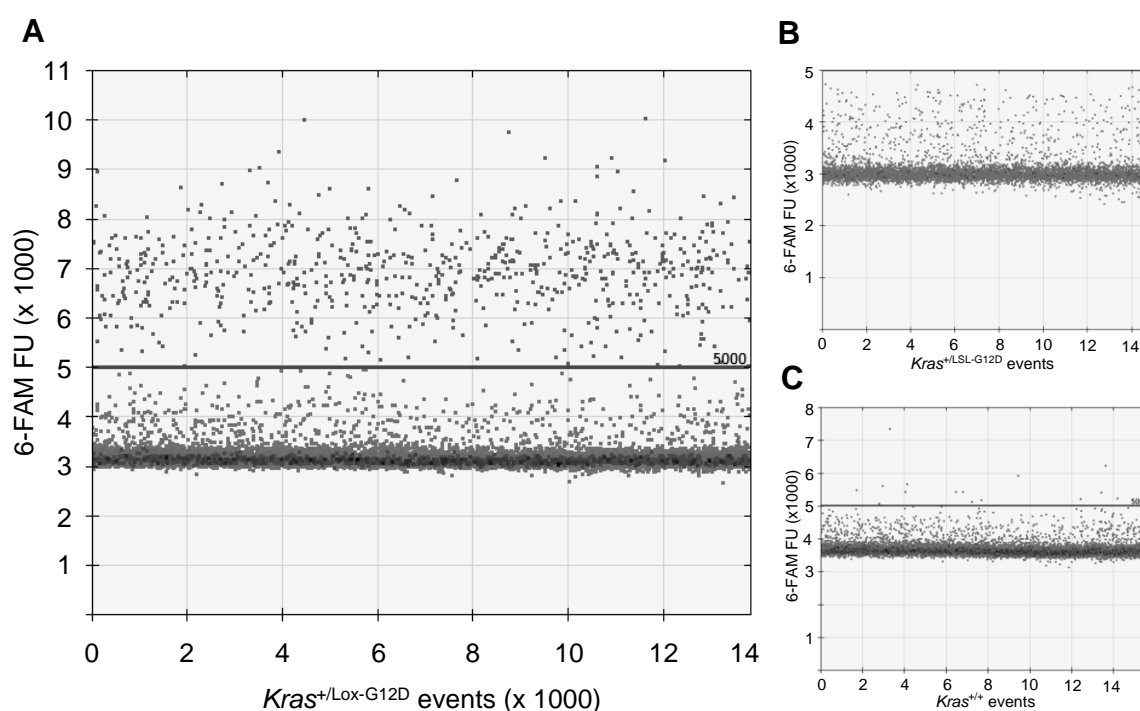


**Figure 3.19 – 1D droplet plots of the 113 bp LNA 1 probe-based assay.** Probe LNA 1 and primers FM 1/RM 2 were run at a $T_a$ of 62.0°C with the following templates: (A) $Kras^{+/Lox-G12D}$ MEF DNA, (B) $Kras^{+/LSL-G12D}$ MEF DNA, (C) $Kras^{+/+}$ DNA. FU, fluorescence units.

Probe LNA 1 was redesigned to eliminate the suspected autofluorescence (Fig 3.20). Since the palindromic nature of LoxP promotes stem-loop formation on both DNA strands, probe LNA 2 was designed to destabilise both stem-loops simultaneously. This 'double destabilisation' strategy was based on a strategy by Esposito *et al.* for improving sequencing through *att*L recombination sites, which are inverted repeat sequences that form stem-loops (Esposito *et al.*, 2003). Probe LNA 2 (Fig 3.20) was based on the sequence of probe LNA 1, except targeting of the 8 bp loop was avoided, because this region is not palindromic and would prevent the probe annealing to both strands. Instead, the probe was targeted only to the right arm of the stem on both strands by reducing the length to 18 nt and shifting the probe further into the *Sal*1 restriction site. Addition of 9 LNA bases to the probe increased the $T_m$ to 70°C, which is higher than the predicted $T_m$ of the amplicon. Since probe LNA 2 was designed to anneal to 5 nt of the 6 nt *Sal*1 restriction site, which is also present in *Kras*[+], there was concern that the probe may have misprimed to the WT sequence. To minimise this possibility, specific bases within the probe that would have formed mismatches with the WT sequence were locked, thus decreasing the stability of base pairing (Fig 3.20).
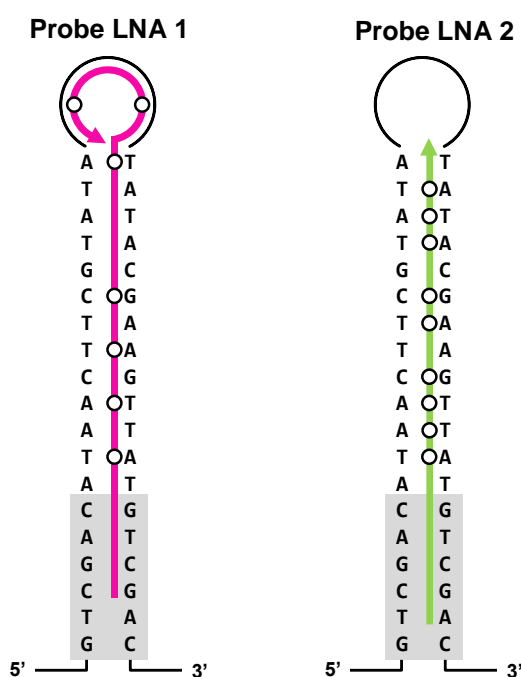


**Figure 3.20 – Targeting of LNA probes 1 and 2 to the recombined LoxP sequence.** Probe LNA 1 targets the loop and right arm of the stem on the reverse strand. Probe LNA 2 targets the right arm of the stem on both strands. Black circles show locked bases. The *Sal*1 sites are in grey.

99

The clustering of positive and negative droplets with probe LNA 2 was determined by performing a $T_a$ gradient as with probe LNA 1, using the same three templates. The optimum $T_a$ was found to be 62.0°C; at this temperature, there was improved discrimination between positive and negative droplets, with negative droplets forming a tight cluster at around 1,000 FU. However, clustering of positive droplets was poor (Fig 3.21 A). No positive droplets were observed with $Kras^{+/LSL-G12D}$ and $Kras^{+/+}$ MEF DNA (Fig 3.21 B, C)
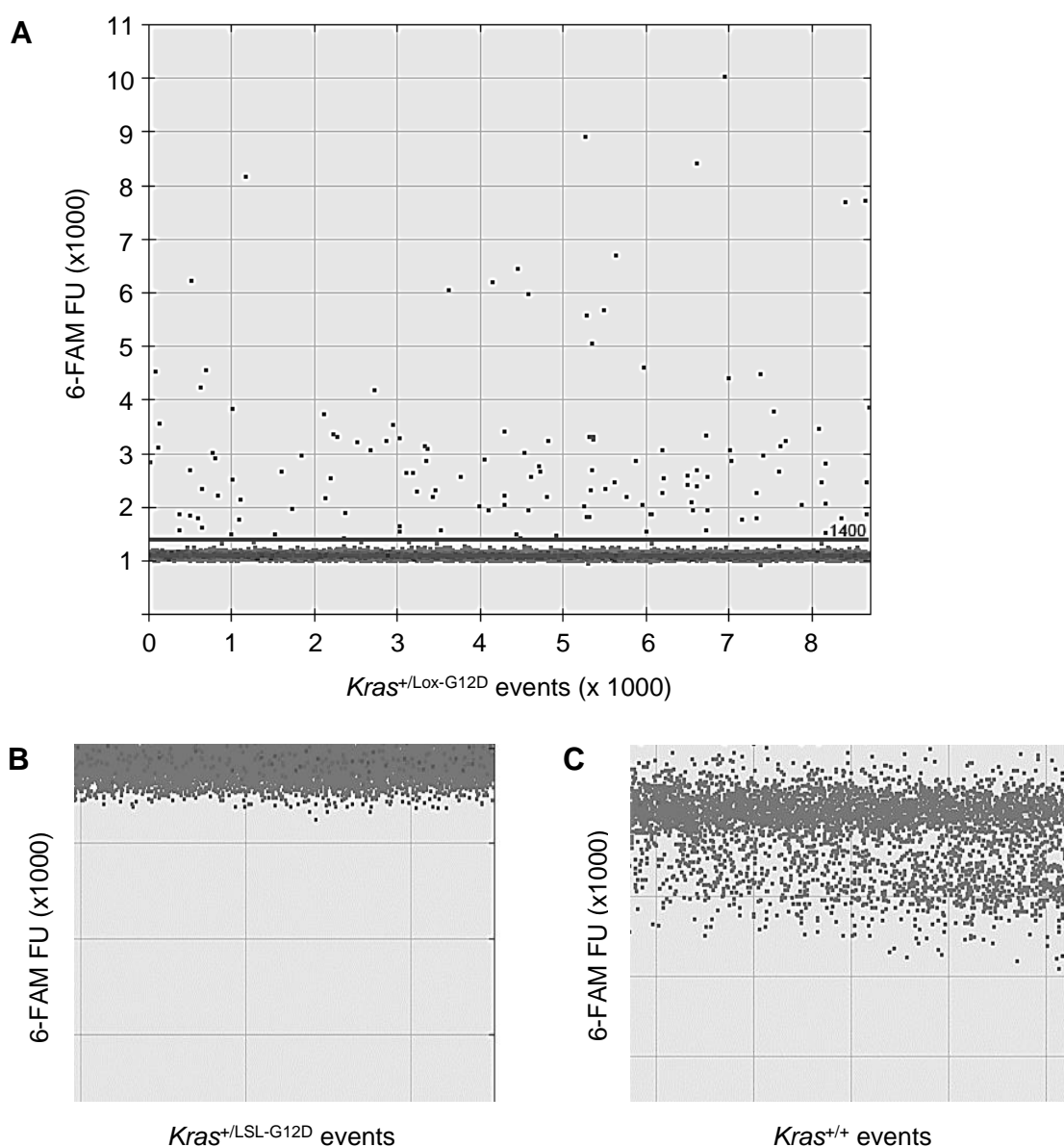


**Figure 3.21 – 1D droplet plots of the 113 bp LNA 2 probe-based assay.** Probe LNA 2 and primers FM 1/RM 2 were run at a $T_a$ of 62.0°C with the following templates: (A) $Kras^{+/Lox-G12D}$ MEF DNA, (B) $Kras^{+/LSL-G12D}$ MEF DNA, (C) $Kras^{+/+}$ DNA. FU, fluorescence units.

Poor droplet clustering, as seen in Figure 3.21, does not permit a confident determination of droplet status (positive/negative) and increases risk of a true positive droplet falling below the threshold. In cfDNA, such positive droplets may signify rare events, and hence accurate determination of droplet status is critical.

To improve droplet clustering and permit setting of a higher threshold, two modifications to the standard thermal cycling protocol were made. First, the cycle number was increased from 40 to 50, in an attempt to increase the fluorescence amplitude of positive droplets. Second, a probe annealing stage was added, theoretically increasing the time for which the LNA probe may anneal to the LoxP sequences in their open conformation.

Modifications were run independently using $Kras^{+/Lox-G12D}$ template. An increase in cycle number from 40 to 50 raised the fluorescent amplitude of each droplet, but gave only a marginal improvement in positive droplet clustering (Fig 3.22 A). Introduction of a probe annealing stage at a range of temperatures (68.0 – 72.0°C) also did not improve clustering (Fig 3.22 B).
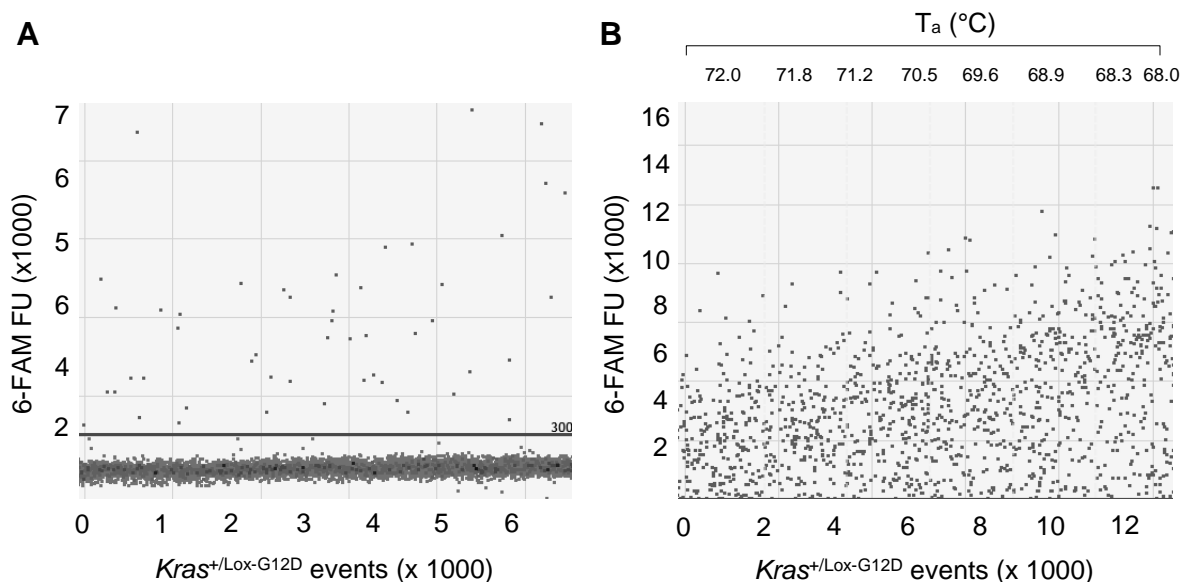


**Figure 3.22 – Cycling protocol optimisation of the 113 bp LNA 2 probe-based assay.** (A) The assay was run for 50 cycles with $Kras^{+/Lox-G12D}$ MEF DNA template. (B) A probe annealing step was introduced to the cycling protocol, with a $T_a$ range of 68.0 – 72.0°C. FU, fluorescence units.

Based on the earlier observation that the amplicon length was critical to the ability of primers flanking the LoxP sequence to amplify (Fig 3.10), the amplicon length of the LNA 2 assay was increased from 113 bp to 133 bp. This theoretically decreased the thermal stability of the LoxP stem-loop at the $T_a$ of the primers, increasing the probability of the LoxP sequence existing in a linear conformation. Probe LNA 2 (Fig 3.20) was tested in assay format with primers FM 1 and R 133 (TCTGCATAGTACGCTATACCCTGTG), which both had a $T_m$ of 65.8°C. A $T_a$ gradient (57.0 – 62.0°C in 1.0°C increments) was carried out with the following DNA templates: *Kras*[+/Lox-G12D] MEF DNA, *Kras*[+/LSL-G12D] MEF DNA and *Kras*[+/+] DNA. The assay was specific to *Kras*[Lox-G12D] and the clustering of positive droplets showed significant improvement from the 113 bp assay, as shown in Fig 3.23.
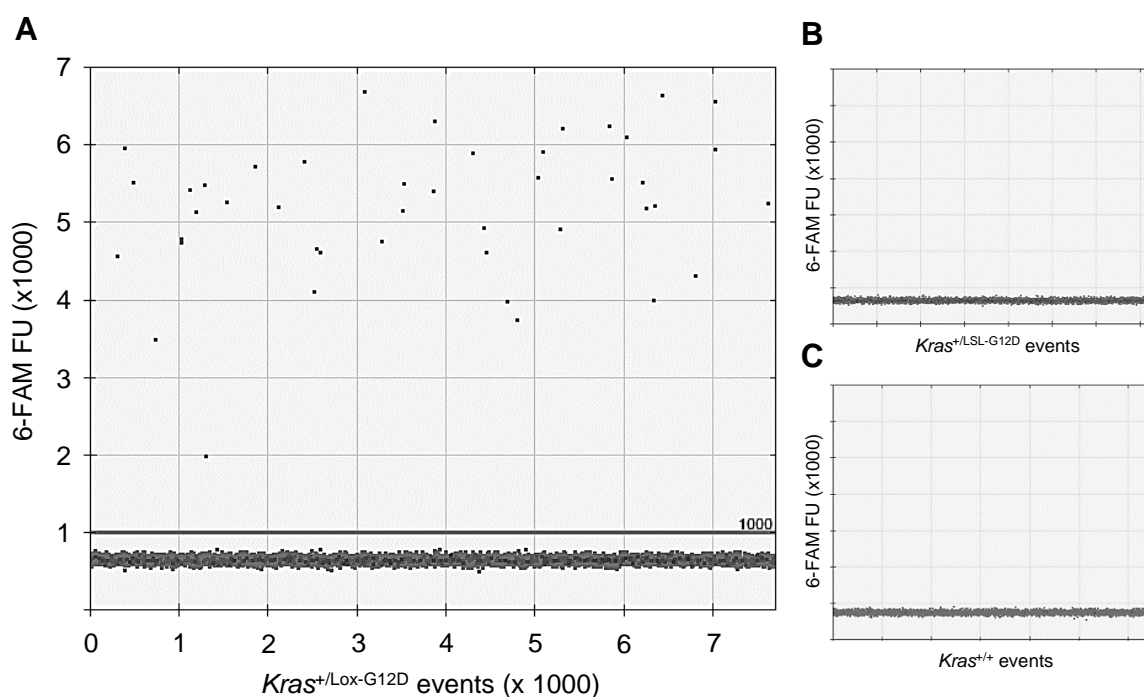


**Figure 3.23 – 1D droplet plots of the 133 bp LNA 2 assay.** Probe LNA 2 and primers FM 1/R 133 were run at a $T_a$ of 62.0°C with the following templates: (A) *Kras*[+/Lox-G12D] MEF DNA, (B) *Kras*[+/LSL-G12D] MEF DNA, (C) *Kras*[+/+] DNA. FU, fluorescence units.

At the earliest stages of tumour progression in the $Kras^{+/LSL-G12D}$ mouse, it is likely that the recombined allele will be absent or exist at a very low frequency in cfDNA, diluted by the unrecombined allele. Therefore, utility of the 133 bp LNA 2 assay in cfDNA requires that detection of $Kras^{Lox-G12D}$ is not compromised by a predominant background of $Kras^{LSL-G12D}$. To determine the sensitivity of the 133 bp LNA 2 assay with ddPCR, $Kras^{+/Lox-G12D}$ MEF DNA was spiked into a background of $Kras^{+/LSL-G12D}$ MEF DNA and serially diluted two-fold over eleven points. The concentration of $Kras^{+/LSL-G12D}$ MEF DNA remained constant in each dilution, and Table 3.3 gives the theoretical copy number and frequency of $Kras^{Lox-G12D}$ in each dilution.

| Dilution | Copy number | Frequency (%) |
|---|---|---|
| 1 | 379 | 33.0 |
| 2 | 189 | 20.0 |
| 3 | 95 | 11.0 |
| 4 | 47 | 5.9 |
| 5 | 24 | 3.0 |
| 6 | 12 | 1.6 |
| 7 | 6 | 0.8 |
| 8 | 3 | 0.4 |
| 9 | 1.5 | 0.2 |
| 10 | 0.75 | 0.1 |
| 11 | 0.38 | 0.05 |

**Table 3.3 – Theoretical copy number and frequency of $Kras^{Lox-G12D}$ in each dilution for sensitivity analysis of the LNA 2 133 bp assay.**

To ensure the concentration of $Kras^{LSL-G12D}$ in each dilution was constant, the LNA 2/FM 1/R 133 assay for recombination was duplexed with the single-locus *Gapdh* 113 bp assay. Copy number of $Kras^{LSL-G12D}$ was calculated by halving the copy number of *Gapdh* (since the transgenic allele is heterozygous for $Kras^{LSL-G12D}/Kras^{Lox-G12D}$). A good correlation between the actual and theoretical copy number was observed down to the tenth dilution, which had a copy number of 0.75 and allele frequency of 0.1% (Fig 3.24). The dropout at the eleventh dilution is consistent with only a 38% chance of a positive droplet being present.
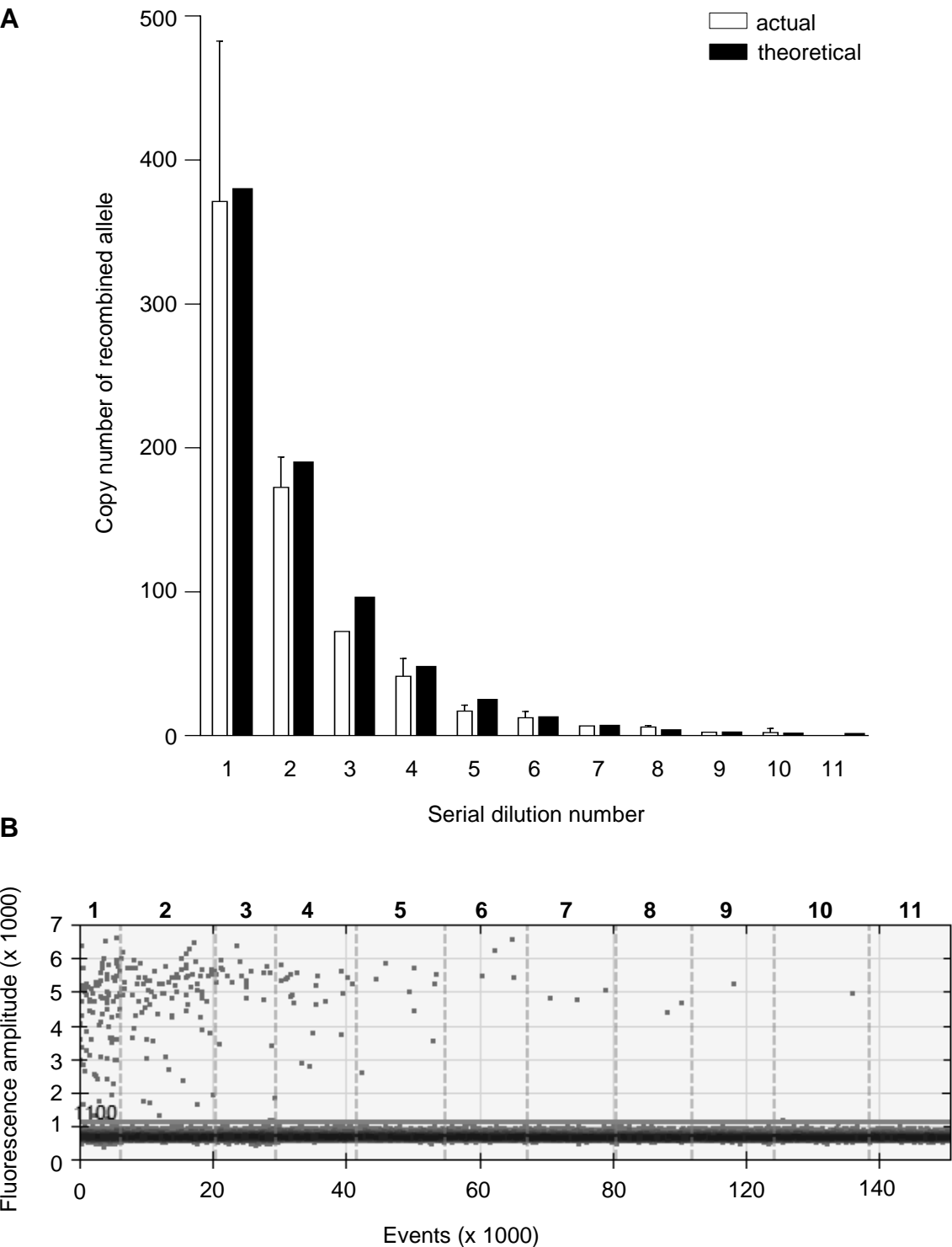
**Figure 3.24 – Experimental determination of LNA 2/FM 1/RM 2 assay sensitivity using ddPCR.** *Kras*+/Lox-G12D MEF DNA was serially diluted into a background of *Kras*+/LSL-G12D MEF DNA. (A) Table to show the theoretical copy number, CN, and frequency, F, of the recombined allele in each serial dilution. (B) A bar chart to show the actual (white) and theoretical (black) CN of the recombined allele in each serial dilution. Error bars represent the mean ± SD of two replicates. (C) 1D droplet plot of experiment, with a manual threshold of 1,100.

### 3.3.5      Final design of assay to detect recombined *Kras*

Through use of LNA technology, specific (Fig 3.23) and sensitive (Fig 3.24) detection of recombined *Kras* in a predominant background of unrecombined *Kras* has been demonstrated by ddPCR.  The final assay targets 133 bp of the recombined *Kras* (*Kras*$^{Lox\text{-}G12D}$) locus and comprises primers that flank the recombined LoxP sequence, along with a dual-labelled (5' 6-FAM™ and 3' BHQ-1™) hydrolysis probe. This probe (Fig 3.25) contains 9 locked bases, which provides strand-invasion properties to the probe and brings the $T_m$ to 70ºC, which, critically, is higher than the predicted $T_m$ of the 133 bp amplicon. Through targeting of the probe to the right arm of the LoxP sequence and *Sal*1 restriction site, the probe can anneal to both template strands and may facilitate 'double destabilisation' of the predicted stem-loop structure formed by the palindromic LoxP sequence on each strand. This can, in turn, promote a linear formation of template for PCR extension. The cycling protocol is shown in Table 3.4 and a schematic of the assay is shown in Figure 3.26.



**6-FAM** – TCGACA[T][A][A][C]T[T][C]G[T][A][T]A – **BHQ-1**

**Figure 3.25 – LNA 2 probe for 'double destabilisation' of LoxP stem-loop structures at the recombined *Kras* locus.**  The probe is 5' 6-FAM™ and 3' BHQ-1™ labelled and contains locked bases, shown in square brackets.

| Stage | Temperature (ºC) | Time | Ramp rate | Cycles |
|---|---|---|---|---|
| **Enzyme activation** | 95 | 10 mins | | 1 |
| **Denaturation** | 95 | 30 secs | 2ºC/sec | 40 |
| **Annealing/extension** | 62 | 30 secs | | 40 |
| **Enzyme deactivation** | 98 | 10 mins | | 1 |

**Table 3.4 – Cycling protocol for LNA2 133 bp assay for detection of *Kras*$^{Lox\text{-}G12D}$.**
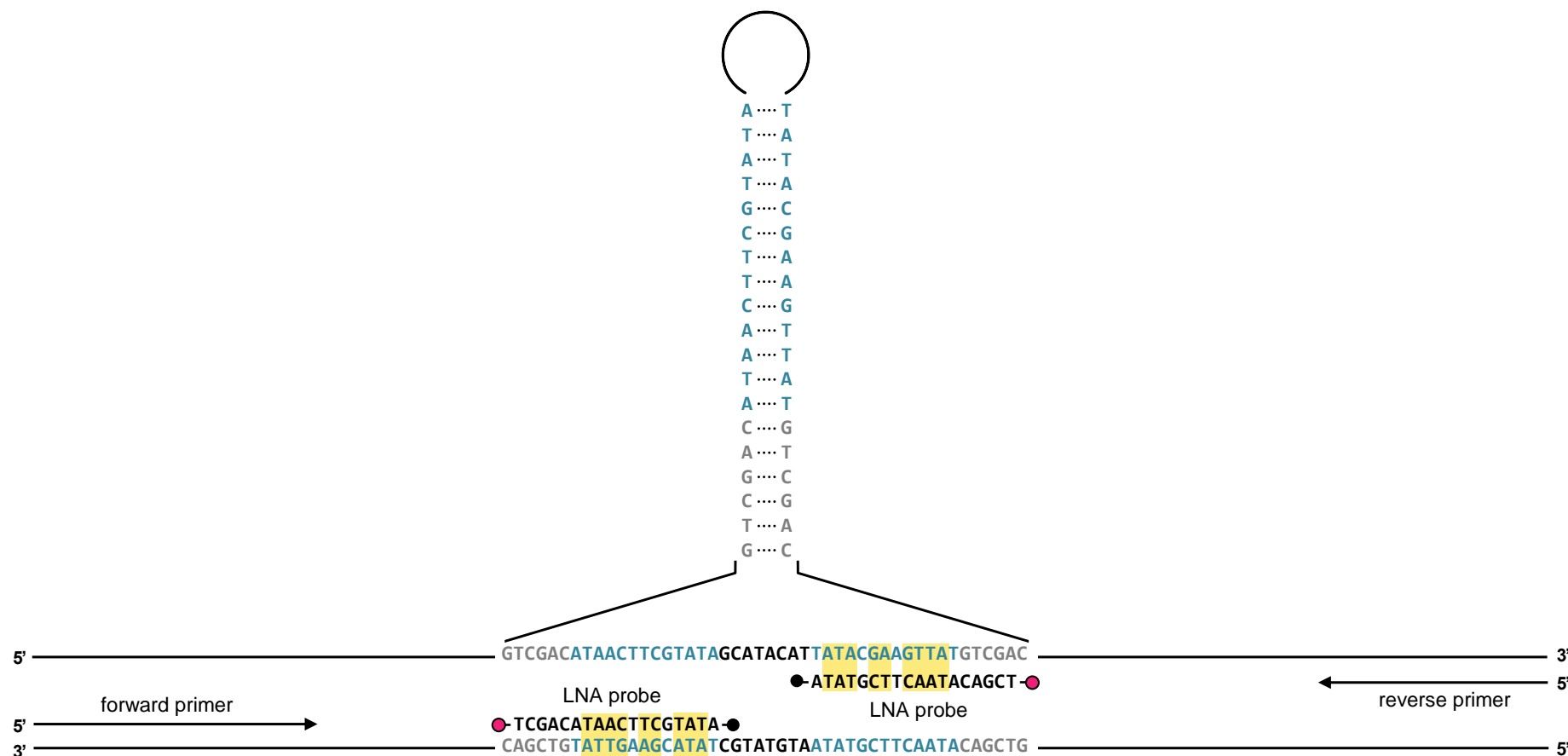
**Figure 3.26 – Schematic of LNA 2 133 bp assay for detection of *Kras*Lox-G12D by ddPCR.** An 18-base FAM-BHQ-1™-labelled hydrolysis probe with 9 locked bases (yellow) targets the right arm of the stem-loop structure on *both* strands, formed by the LoxP sequence. Unlabelled primers flank the recombined LoxP sequence, producing an amplicon of 133 bp. The stem-loop on the reverse strand is not shown.

### 3.3.6    Summary of assays developed for analysis of cfDNA

In this Chapter, PCR assays were designed for total quantification of cfDNA (targeting either single-locus *Gapdh* or 44 *Gapdh* pseudogenes) and sensitive detection of unrecombined and recombined $Kras^{G12D}$ alleles in cfDNA.

Quantification of total cfDNA over time may permit correlation with tumour burden assessed via computed tomography (CT) imaging (Haines *et al.*, 2009; Namati *et al.*, 2010a). Combined use of the assays to detect $Kras^{LSL\text{-}G12D}$ and $Kras^{Lox\text{-}G12D}$ will permit the calculation of a recombination frequency with respect to the transgenic allele, and monitoring dynamic changes in this parameter over time will enable a comparison of tumour burden and ctDNA level.

### 3.3.7      Assessment of sequence variation in recombined LoxP

Mutational analysis of LoxP sequences before and after Cre-mediated recombination has not been reported previously. Since the assays developed in this chapter to detect Cre-Lox recombination require annealing of a probe to the LoxP sequence, mutations in and around LoxP may compromise assay performance. To investigate possible sequence variation, DNA was extracted from an FFPE core biopsy of Kras[+/Lox-G12D] mouse lung tissue, which was confirmed to contain multiple adenomas by histological analysis with H/E (Fig 3.27). A 180 bp region surrounding the recombined LoxP sequence was PCR-amplified with the 140/180 bp primers detailed in Table 2.4 and cloned into a vector. Clones were screened by PCR and 40 clones containing the 180 bp insert were sequenced.



500 µm

5x
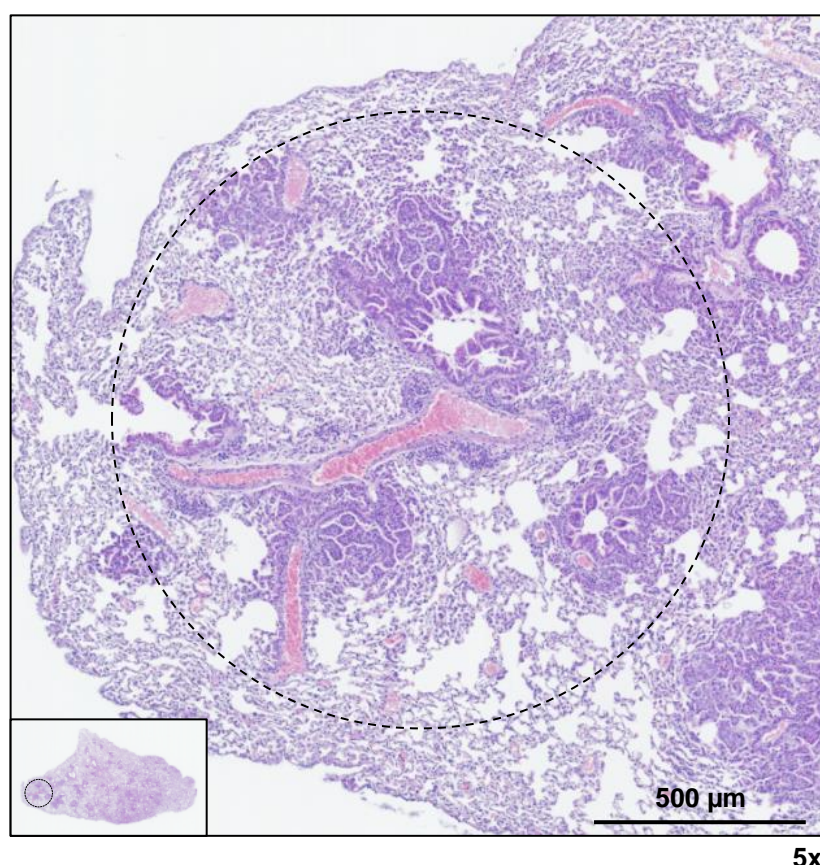
**Figure 3.27 – Location of core biopsy from Kras[+/Lox-G12D] mouse lung**. A 1.5 mm diameter core biopsy was punched from a region of Kras[+/Lox-G12D] mouse lung tissue, which was confirmed to contain multiple adenomas by H/E analysis.

**Figure 3.28 – Alignment of reads for sequences within and around the recombined LoxP sequence in *Kras*[Lox-G12D]**. 140 bp/180 bp primers flanking the recombined LoxP sequence were used to amplify a 180 bp sequence in *Kras*[Lox-G12D]. Mutations are shown in grey. Primer sequences are not shown.

Sequence alignment revealed point mutations within 12 of the 40 clones containing recombined LoxP (Fig 3.28). These mutations clustered within and around the *Sal*1 restriction sites, and also occurred in intronic sequence, as shown in Figure 3.28. 11 of the 12 mutations were transitions (4 A>G; 1 G>A; 4 T>C; 2 C>T), and all mutations were confirmed by sequencing the opposite strand. To confirm these mutations were not induced by errors in DNA polymerase during PCR amplification of the insert sequence, the 140 bp WT region of *Kras* (co-amplified by the same primers) was also cloned. 40 clones were sequenced and no mutations were identified, thus confirming the mutations observed following Cre-mediated recombination of LoxP sequences were induced by the recombination process.

As shown in Figure 3.29, the presence of point mutations within and around the recombined LoxP sequence may prevent annealing of primers and probe in the ddPCR for detection of recombined *Kras* (Fig 3.26).
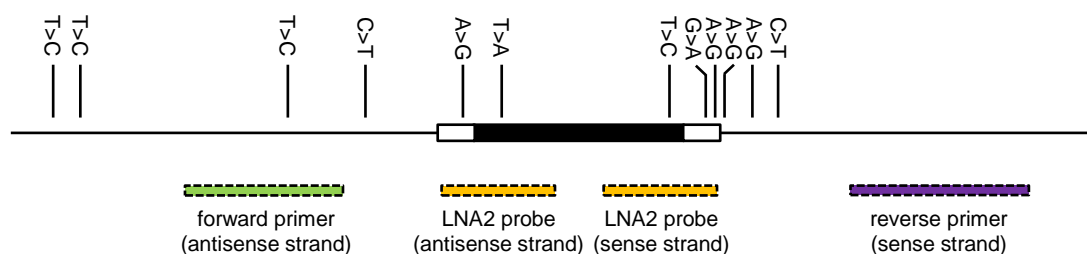


**Figure 3.29 – Cre-Lox recombination induces mutations within and around the recombined LoxP sequence.** A 180 bp region surrounding the recombined LoxP sequence (black; flanked with Sal1 restriction sequences, white) was PCR-amplified from DNA isolated from a Kras[+/Lox-G12D] mouse lung. The region was cloned and 40 clones were isolated for sequencing. Vertical lines indicate the positions of the mutations. Annealing positions of oligonucleotides in the recombined allele assay are shown in dashed boxes.

## 3.4 Discussion

The motivation for the development of quantitative PCR strategies to analyse cfDNA in the Kras$^{+/LSL-G12D}$ mouse was the result of a proof-of-concept study presented in Chapter 4, in which the recombined *Kras* allele (*Kras*$^{Lox-G12D}$) was identified in the cfDNA of Kras$^{+/LSL-G12D}$ mice 12 weeks post-adenoviral Cre infection, using end-point PCR. Prior to this work, evidence of ctDNA was first demonstrated in the Braf$^{+/LSL-V600E}$ model of lung cancer (Chapter 4), but given the relative infrequency of *Braf* mutations in human lung cancer, the *Kras*$^{G12D}$-driven model was deemed more clinically relevant, and the Braf$^{+/LSL-V600E}$ model was not evaluated further. Interestingly, detection of ctDNA in both mouse models preceded the development of overt cancer; instead, mice exhibited a spectrum of preneoplastic lesions that closely resembled those of the cognate human disease. This finding raised the prospect of using conditional mouse models to investigate ctDNA as a disease biomarker in various settings, with a particular emphasis on early detection.

### 3.4.1 Quantification of total cfDNA

In human patients with cancer, cfDNA is a dynamic biomarker, rising in levels as the disease progresses and regressing to baseline levels following curative resection (Swystun *et al.*, 2011; To *et al.*, 2003). This phenomenon is also well reported in xenograft mouse models (Rago *et al.*, 2007; Thierry *et al.*, 2010; Mouliere *et al.*, 2011; Mouliere *et al.*, 2013; Czeiger *et al.*, 2011; Gorges *et al.*, 2012). Therefore, total cfDNA levels may be informative in GEM models of cancer. cfDNA was quantified in Kras$^{+/LSL-G12D}$ mice using a qPCR assay targeting 113 bp of *Gapdh*. The observation that cfDNA levels were significantly elevated in tumour-bearing mice relative to wild-type mice at end-point provided further evidence of ctDNA in the Kras$^{+/LSL-G12D}$ mouse, although cfDNA levels did not correlate with tumour burden, as determined by histological analysis. One possible explanation is that tumour burden was estimated based on a single tissue section, which may have poorly represented the overall tumour burden.

111

Also, end-point cfDNA levels varied widely among tumour-bearing mice, likely due to variable tumour burden; inconsistent delivery of adenoviral Cre to the lungs is major limitation of nasal delivery in conditional models (DuPage et al., 2009).

An additional qPCR assay targeting 44 *Gapdh* pseudogenes was developed for a more sensitive quantification of total cfDNA, such as in small blood samples taken for longitudinal analysis (Chapter 4.3.3). This assay showed linearity down to 2 pg/reaction, which is less than a genome equivalent (3.3 pg).

### 3.4.2      Detection of *Kras*$^{LSL-G12D}$

A qPCR assay for the detection of *Kras*$^{LSL-G12D}$ was developed to enable the recombination frequency of the transgenic allele (along with an assay to detect *Kras*$^{Lox-G12D}$) to be calculated in cfDNA. A simple MGB probe strategy based on the Weis *et al.* assay was devised, in which a hydrolysis probe was targeted to the left unrecombined LoxP sequence, and flanking primers were positioned within intron 1 and the STOP sequence of the LSL element. The assay showed a minimum linear dynamic range of 160 – 10,000 pg/reaction and an efficiency of 101%, as determined by standard curve analysis.

### 3.4.3      Detection of *Kras*$^{Lox-G12D}$

Initial demonstration of ctDNA in the *Braf*$^{V600E}$- and *Kras*$^{G12D}$-driven lung cancer models was achieved with end-point PCR, generating amplicons corresponding to the recombined alleles that were easily resolved electrophoretically. However, several pitfalls of end-point PCR for analysis of ctDNA were evident. For detection of the *Kras*$^{Lox-G12D}$ allele, the band corresponding to the recombined allele was faint, even with heterozygous cell line (*Kras*$^{+/Lox-G12D}$) template, which should have produced two bands of approximately equal intensity, corresponding to *Kras*$^+$ (140 bp) and *Kras*$^{Lox-G12D}$ (180 bp). This suggests inefficient amplification of *Kras*$^{Lox-G12D}$ relative to *Kras*$^+$, possibility in part due to differences in amplicon length. Also, end-point PCR produces only semi-quantitative data since measurements are made at the plateau phase of amplification.

To improve the sensitivity of *Kras*[Lox-G12D] detection and permit a quantitative analysis of cfDNA in the Kras[+/LSL-G12D] mouse, assays based on qPCR and ddPCR were developed. Initial qPCR strategies were based on SYGR Green analysis and avoided co-amplification of *Kras*[+] by positioning one or both primers within/overlapping into the recombined LoxP sequence. These strategies failed to produce PCR products at the desired size, instead favouring non-specific products and primer dimers. It was hypothesised at this stage that, following PCR denaturation, the palindromic nature of the LoxP sequence may promote the formation of a stem-loop structure, which may have a high $T_m$ and thus occlude access to primers and DNA polymerase at a $T_a$ of approximately 60°C.

*In silico* prediction of secondary structure using mFold software predicted a stem-loop structure with a $T_m$ of approximately 55.6°C. However, since the LSL element was cloned into *Kras*[+] via a 6 bp *Sal*1 restriction site during generation of the transgene, the recombined LoxP sequence is flanked by these palindromic restriction sites, resulting in a significantly longer (19 bp) stem-loop structure with a predicted $T_m$ of approximately 65.5°C. Thus, in the initial experiments with conventional length primers, it is likely that some primers could not anneal because their target sequences were occluded by the stable stem-loop structure.

An alternative approach was attempted in which primers were lengthened to increase the $T_m$ above the $T_m$ of the predicted stem-loop structure of recombined LoxP. Theoretically, the recombined LoxP sequence would exist in a single-stranded conformation at the $T_a$ of the primers, permitting primer access. However, this PCR strategy also failed to produce products at the expected size.

Based on a published qPCR strategy by Weis *et al.* (2010) for targeting the recombined LoxP sequence in a *Dnmt3a*-deleted locus, high-$T_m$ TaqMan MGB probes were designed to target the recombined LoxP sequence directly (Weis *et al.*, 2010). Such probes would, in theory, destabilise or prevent formation of the stem-loop structure to allow primers to anneal at a $T_a$ of approximately 60°C. Three TaqMan MGB probes were designed to destabilise different regions of the stem-loop, including the left arm and loop region, right arm and loop region, and left arm alone. Each probe was tested with all primer pairs that produced

products, but no amplification signal was observed from any primer/probe combination.

The reason for failure of the TaqMan MGB probe strategy to permit amplification across the recombined LoxP sequence was unclear. A paper published by Hertoghs *et al.* reported the use of LNA modification technology to produce hydrolysis probes that could invade double-stranded plasmid DNA and displace non-complementary strands (Hertoghs *et al.*, 2003). It was hypothesised that the 'strand invasion' property of LNA nucleotides, as reported in this paper, could provide further destabilisation potential to the MGB probes used previously. To this end, one MGB probe was redesigned to incorporate 7 locked nucleotides, which maintained the $T_m$ at 70°C despite removing the MGB moiety. The modified probe was tested with primers flanking the recombined LoxP sequence and was successful in producing an amplification signal. However, a standard curve of *Kras*$^{Lox-G12D}$ DNA showed the assay had poor efficiency, and a subsequent experiment in which *Kras*$^{Lox-G12D}$ DNA was serially diluted into a constant background of *Kras*$^{LSL-G12D}$ DNA revealed that the assay was sensitive only down to a 6.25% frequency of *Kras*$^{Lox-G12D}$, which is unacceptable for use in cfDNA where the ctDNA fraction may be very low (Schwarzenbach *et al.*, 2008). This suggested that *Kras*$^{LSL-G12D}$ was outcompeting *Kras*$^{Lox-G12D}$ for the probe.

To overcome the suppressive effect of *Kras*$^{LSL-G12D}$ on amplification of *Kras*$^{Lox-G12D}$, the assay was translated to ddPCR methodology. Partitioning of template into many droplets would theoretically prevent the two alleles existing in the same droplet, providing sufficient dilution of template. Indeed, detection of *Kras*$^{Lox-G12D}$ was unaffected by the presence of *Kras*$^{LSL-G12D}$ and *Kras*$^+$, the latter of which competes with *Kras*$^{Lox-G12D}$ for both primers. Droplet discrimination was poor, with the majority of positive droplets clustering loosely at low fluorescence amplitude, rendering discrimination between positive and negative droplets difficult. A clustering of droplets towards baseline was observed with all templates and the NTC, suggesting autofluorescence from the FAM fluorophore.

The probe was redesigned to eliminate autofluorescence and the sequence was modified such that the new probe could anneal to *both* DNA strands, providing

destabilisation to both stem-loop structures simultaneously. This strategy was based on an oligonucleotide blocking strategy reported by Esposito *et al.* for improving sequencing through *att*L recombination sites, which are inverted repeat sequences that are prone to stem-loop formation (Esposito *et al.*, 2003).

Whilst the new probe did not exhibit autofluorescence, clustering of positive droplets was poor and the majority of droplets were of low fluorescence, again making discrimination between positive and negative droplets difficult. Poor clustering of positive droplets may be explained by the Monte-Carlo effect, in that there were three probabilistic events associated with primer/probe annealing:

(i)      existence of *both* LoxP sequences in a linear conformation at the $T_a$ of the probe;

(ii)     annealing of the probe to the linear LoxP sequen<u>ce</u>s at this $T_a$; and

(iii)    annealing of each primer to a linear LoxP sequence at the primer $T_a$.

Each of these events can influence the overall efficiency of amplification and, in turn, determine the fluorescence amplitude of each droplet.

An increase in cycle number and incorporation of a probe annealing step did not significantly improve clustering, but a 20 bp increase in amplicon length had a dramatic effect, with positive droplets clustering at higher fluorescence amplitude, permitting more confident thresholding. The assay was not further optimised at this point. The analytical sensitivity of the final assay for $Kras^{Lox-G12D}$ detection was evaluated by serially diluting $Kras^{Lox-G12D}$ DNA into $Kras^{LSL-G12D}$ DNA. The assay was sensitive down to the tenth dilution, corresponding to a copy number of 0.75 and an allele frequency of 0.1%, which is suitable for cfDNA analysis.

After the assays in this Chapter were developed, the possibility of mutations at low frequency in the recombined LoxP sequence was raised. Such mutations could impair the performance of the PCR assays developed in this chapter to detect Cre-Lox recombination, since annealing of oligonucleotides to their target sequences may be compromised by base mismatches. Cloning of the genomic region around the recombined LoxP sequence in FFPE tissue from a $Kras^{+/LSL-G12D}$ mouse confirmed this observation, with 30% of clones showing a point

mutation in this region. Mutations were present in both the LoxP sequence and flanking regions (Sal1 sites and intronic sequences). Of particular concern were mutations within the LoxP sequence, since the assay for detection of *Kras*[Lox-G12D] requires annealing of an LNA-modified probe to the LoxP sequence on both strands. Moreover, considering the mismatch discrimination properties of LNA bases, it is likely that single-base mismatches would abrogate annealing of the probe (Owczarzy *et al.*, 2011).

## 3.5      Conclusions

A PCR strategy for quantitative detection of the recombined *Kras* allele in cfDNA of the Kras[+/LSL-G12D] mouse model was successfully designed. The propensity of the palindromic LoxP sequence to fold into a stable stem-loop during PCR presented a major challenge in the design of a PCR strategy that would provide sufficient destabilisation of this structure. Amplifying across the recombined LoxP sequence was particularly challenging because its flanking *Sal*1 restriction sites extended the stem-loop by another 6 bp to 19 bp. The shorter (16 bp) stem-loop formed by the unrecombined LoxP sequences were easily and efficiently destabilised using a conventional TaqMan MGB probe strategy based on the assay reported by Weis *et al.* (Weis *et al.*, 2010). However, destabilising the stem-loop formed by the recombined LoxP sequence was significantly more problematic, requiring an LNA-modified probe strategy to promote strand invasion. However, an assay with specificity for the recombined *Kras* allele and a low limit of detection was developed successfully. The LNA strategy may be applicable to other conditional mouse models in which the recombined LoxP sequence forms thermostable stem-loops that cannot be destabilised using conventional MGB probe-based strategies. The assays developed in this Chapter permitted quantitative analysis of cfDNA in further cohorts of Kras[+/LSL-G12D] mice, presented in Chapter 4.

# Chapter 4

# Molecular analysis of cfDNA in the Kras$^{+/LSL-G12D}$ mouse model of lung adenocarcinoma development

## 4.1        Introduction

### 4.1.1        Background

cfDNA has been investigated in late-stage human NSCLC, where molecular alterations such as mutations, copy number variations, structural rearrangements and epigenetic patterns have been shown to mirror those of primary and metastatic tumours (Elshimali *et al.*, 2013). However, few studies have investigated the molecular features of cfDNA in the early and preneoplastic stages of the disease. Studies of this kind are hampered by the challenge of identifying individuals with early, preneoplastic lesions that may progress to malignancy. Instead, owing to the lack of cost-effective screening initiatives and the relatively late presentation of the disease, two-thirds of NSCLC patients present with metastatic spread that is refractory to curative treatment (Molina *et al.*, 2008).

Clinically-relevant mouse models could provide valuable insights into the utility of cfDNA as an early diagnostic biomarker, for several reasons. First, cfDNA analysis in humans is challenging due to intrinsic genetic and environmental heterogeneity between patients. Experimental mice are generated from inbred strains and can be housed in identical environmental conditions, thus minimising such extraneous variability. Second, some mouse models can recapitulate the initiating stages of tumour development in humans, which are virtually undetectable in asymptomatic patients using current screening methods. In addition, the ability to control disease onset permits the acquisition of blood samples at defined time points, before and during the course of tumour progression. Matched tumour tissue can be harvested to determine the stage of disease at the time of blood collection, and small animal imaging technologies may be used to non-invasively follow tumour progression longitudinally (Namati *et al.*, 2010b). Importantly, age- and sex-matched wild type littermates can be compared to experimental mice in order to confirm disease-specific changes in cfDNA.

## 4.2      Aims and objectives

The aims of this chapter were to determine:

i)      whether cfDNA levels increase over time and correlate with the stage of disease in the Kras[+/LSL-G12D] model; and

ii)     the earliest stage of disease at which ctDNA, i.e. the recombined (*Kras*[Lox-G12D]) allele, can be detected.

The objectives were to:

i)      quantify total cfDNA at longitudinal time points to correlate with tumour burden, assessed by µCT imaging;

ii)     calculate *Kras*[Lox-G12D] allele frequencies in cfDNA at end-point and correlate with tumour burden, assessed by µCT imaging; and

iii)    track back to previous time points to determine the minimum tumour burden at which the *Kras*[Lox-G12D] allele can be detected.

## 4.3        Results

### 4.3.1        Proof-of-concept work

#### 4.3.1.1        Braf[+LSL-V600E] model

The Braf[+/LSL-V600E] model of lung adenocarcinoma is a clinically relevant model that exhibits a range of preneoplastic lesions (Chapter 1.7.1) (Dankort *et al.*, 2007). To determine whether this model is a suitable model for analysis of cfDNA, a possible tumoural component of cfDNA was investigated using end-point PCR and agarose gel analysis.

Since this model was generated by the Pritchard lab, the *Braf*[LSL-V600E] allele is well characterised and the sequences surrounding the LoxP sequences are known. A pair of primers flanking the three LoxP sequences were used to assess for complete Cre-mediated recombination in cfDNA. These primers were obtained from the Pritchard lab, where they are used for animal genotyping purposes. The forward primer is targeted to intron 14 and the reverse primer is targeted to a short sequence that was used to clone-in in the right LoxP sequence. In this configuration, the primers are specific to *Braf*[Lox-V600E], producing an amplicon of 205 bp, and do not amplify *Braf*[LSL-V600E] due to the long LSL element of several kb in length (Fig 4.1).
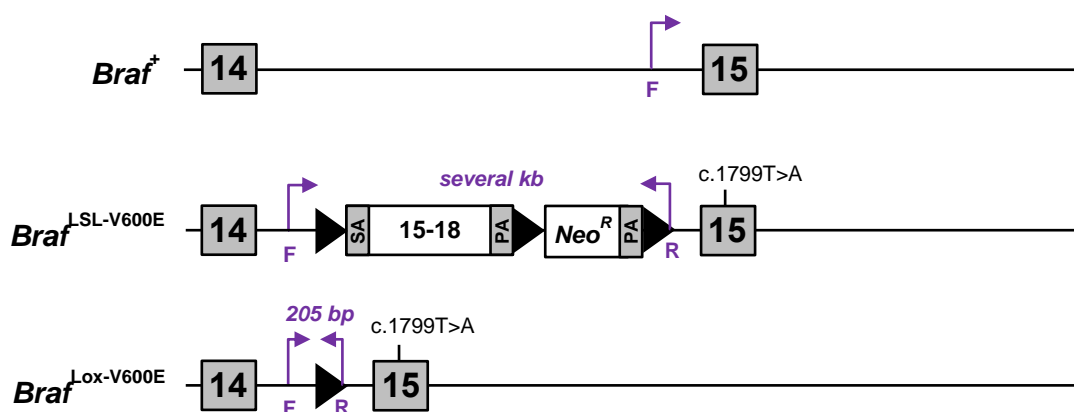


**Figure 4.1 – Overview of the end-point PCR strategy for detection of *Braf*[Lox-V600E] in cfDNA.** *Braf*[Lox-V600E] is detected by flanking the LSL element with forward and reverse primers. The reverse primer does not anneal to *Braf*[+] because it targets a unique sequence used to clone=in the right LoxP sequence. Primers amplify a 205 bp region of *Braf*[Lox-V600E]. *Braf*[LSL-V600E] does not amplify because the STOP sequence within the LSL element cannot be fully extended during the PCR extension step.

45 µL cfDNA (entire eluate) was first analysed from cardiac blood in a cohort of 4 mice (3 Braf[+/LSL-V600E] and 1 Braf[+/+]; Cohort 1) at 12 weeks following nasal administration of 1 x 10$^8$ PFU adenoviral Cre. At this time, all lungs exhibited alveolar hyperplasia throughout, bronchial epithelial hyperplasia and extensive well-circumscribed benign adenomas with bronchiolar involvement, as assessed by histological analysis with haematoxylin and eosin (H/E) stain. Fig 4.2 is a representative example of a Cohort 1 Braf[+/Lox-V600E] mouse lung at end-point. 3/3 cfDNA samples from Braf[+/LSL-V600E] mice showed evidence of *Braf*[Lox-V600E] (Fig 4.3) by end-point PCR using primers as positioned in Fig 3.4.

Since Cre expression was driven by the high-level constitutive CMV promoter, adenoviral Cre had potential to infect and induce recombination in any cell type, including blood cells if the virus crossed the respiratory membrane. To rule out direct recombination in the blood, genomic DNA was extracted from circulating leukocytes and subjected to PCR as in Figure 4.3. There was no evidence of leukocyte recombination in any mice (data not shown).



500 µm

5x
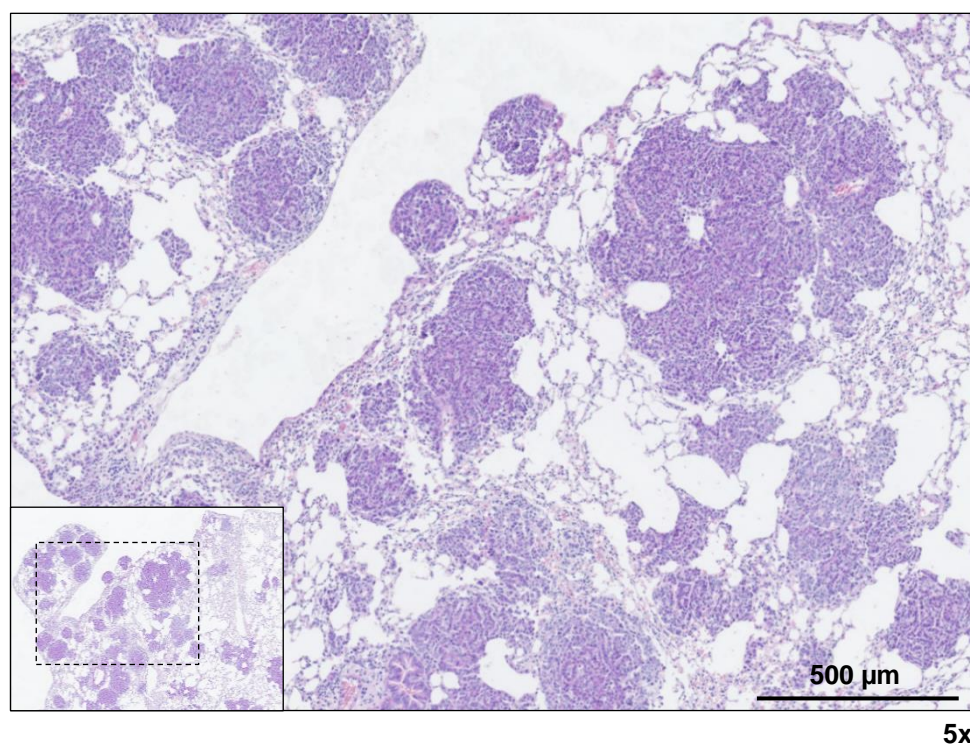
**Figure 4.2 – Representative H/E section of *Braf*[+/Lox-V600E] lung tissue 12 weeks post-administration of 1 x 10$^8$ PFU adenoviral Cre.** After 12 weeks of tumour development, lungs exhibited hyperplasia throughout, bronchial epithelial hyperplasia and extensive well-circumscribed adenomas with bronchiolar involvement.
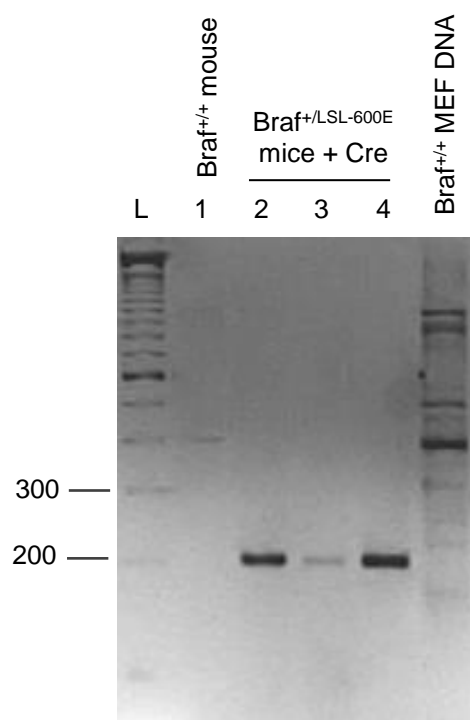
**Figure 4.3 – *Braf*[Lox-V600E] can be detected in cfDNA at end-point.** Using primers that flank the LSL element in Braf[LSL-V600E], the *Braf*[Lox-V600E] allele (205 bp) was analysed by PCR in cfDNA from 3 *Braf*[+/LSL-V600E] mice and 1 *Braf*[+/+] mouse after 12 weeks of adenoviral Cre administration at 1 x 10[8] PFU. MEF DNA carrying WT alleles (*Braf*[+/+]) was also included. L, ladder.

These results were encouraging because they demonstrated, for the first time, evidence of tumour-derived DNA (i.e. the recombined *Braf* allele, *Braf*[+/Lox-V600E]) in terminal blood of genetically engineered mice with preneoplastic lesions and no evidence of overt cancer.

### 4.3.1.2   Kras[+LSL-G12D] model

After successfully detecting the recombined *Braf* allele in cfDNA of Braf[+/LSL-V600E] mice at end-point, a similar approach was taken to investigate a possible tumour-derived component of cfDNA in the Kras[+/LSL-G12D] model (Chapter 1.7.2) (Jackson *et al.*, 2001). The approach used end-point PCR and agarose gel analysis as before.

At the start of the project, the genomic position of the Lox-STOP-Lox (LSL) element in *Kras*[+/LSL-G12D] was not published, and this information was critical to the design of a short-amplicon qPCR assay for Cre-Lox recombination. To identify the approximate insert position, genotyping primers that flanked the recombined LoxP sequence (producing 425 bp and 465 bp amplicons for *Kras*[+] and *Kras*[Lox-G12D] alleles, respectively) were used as a starting point, and the genomic region was gradually shortened using a 'walking in' PCR approach with multiple primer pairs, until amplicon lengths of 140 bp for *Kras*[+] and 180 bp for *Kras*[Lox-G12D] were achieved using a single primer pair flanking the LSL element. The 5,646 bp STOP sequence within the LSL element prevented amplification of *Kras*[LSL-G12D] (Fig 4.4). Thus, the primer pair was suitable to assess for Cre-Lox recombination in cfDNA.
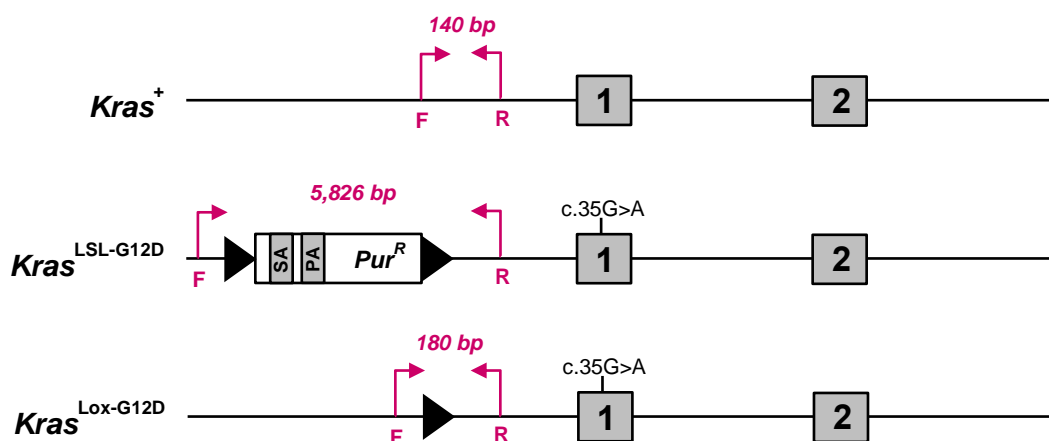


**Figure 4.4 – Overview of the end-point PCR strategy for detection of *Kras*[Lox-G12D] in cfDNA.** *Kras*[Lox-G12D] is detected by flanking the LSL element with forward and reverse primers. Primers amplify a 140 bp region of *Kras*[+] and a 180 bp region of *Kras*[Lox-G12D] (WT + 40 bp LoxP). *Kras*[LSL-G12D] does not amplify because the 5,646 bp STOP sequence within the LSL element cannot be fully extended during the PCR extension step.

45 µL cfDNA was first analysed from cardiac blood in a cohort of 7 mice (Cohort 1) at 12 weeks following nasal administration of $1 \times 10^8$ PFU adenoviral Cre. At this time, all lungs exhibited widespread alveolar hyperplasia, bronchial epithelial hyperplasia and extensive benign adenomas with bronchiolar involvement, as assessed by histological analysis H/E stain. Fig 4.5 is a representative example of a Cohort 1 *Kras*[+/Lox-G12D] lung at end-point. 3/7 cfDNA samples showed evidence of *Kras*[Lox-G12D] (Fig 4.6) by end-point PCR using primers as positioned in Fig 3.6.

As with the Braf[+/LSL-V600E] mice, genomic DNA was extracted from circulating leukocytes and subjected to PCR as in Figure 4.6. There was no evidence of leukocyte recombination in any mice (data not shown).
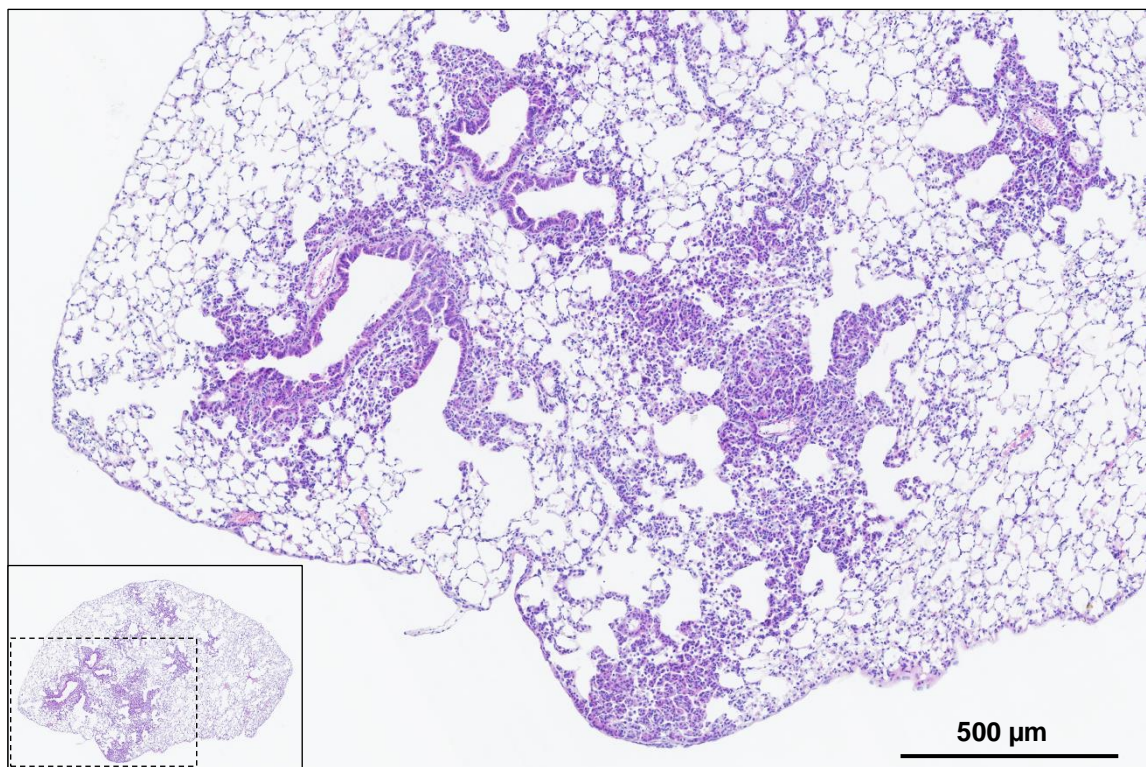


5x

**Figure 4.5 – Representative H/E section of *Kras*[+/Lox-G12D] lung tissue 12 weeks post-administration of $1 \times 10^8$ PFU adenoviral Cre.** After 12 weeks of tumour development, lungs exhibited widespread hyperplasia, bronchial epithelial hyperplasia and extensive adenomas with bronchiolar involvement.
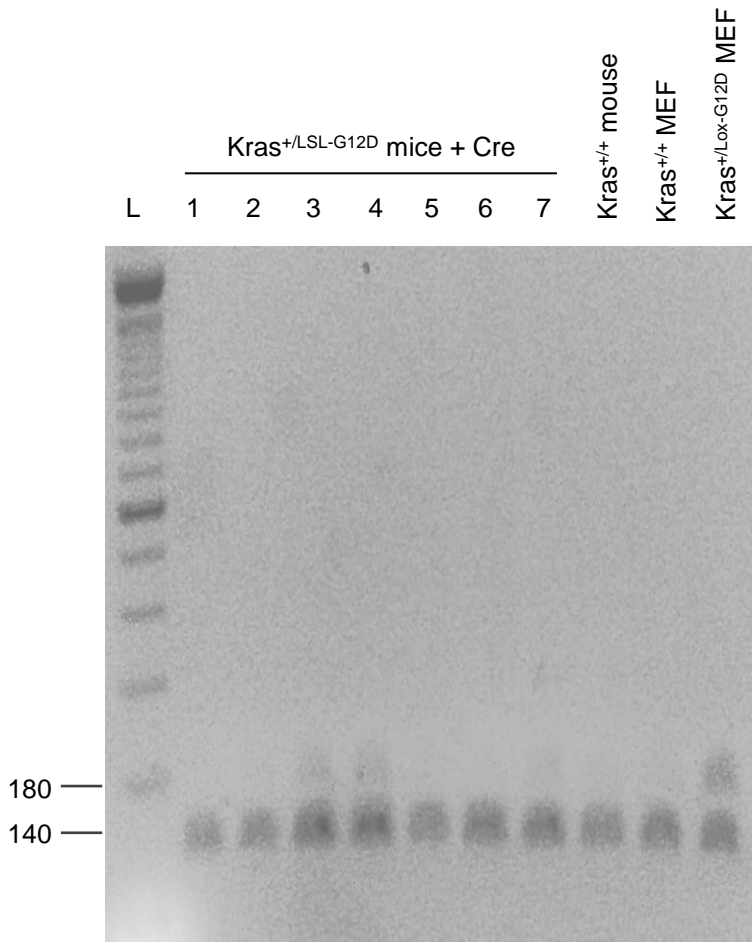
**Figure 4.6 – *Kras*Lox-G12D can be detected in cfDNA at end-point.** Using primers that flank the LSL element in *Kras*LSL-G12D, the *Kras*Lox-G12D (180 bp) was analysed by PCR in cfDNA from 7 *Kras*+/LSL-G12D mice and 1 *Kras*+/+ mouse after 12 weeks of adenoviral Cre administration at 1 x $10^8$ PFU. MEF DNA carrying *Kras*Lox-G12D or *Kras*LSL-G12D alleles was used as positive and negative controls, respectively. L, ladder.

These results were encouraging because they demonstrate, for the first time, that the recombined *Kras* allele can be detected in cfDNA isolated from terminal blood of *Kras*+/LSL-G12D mice, despite no evidence of overt lung adenocarcinoma. The next step was to quantitate levels of cfDNA at end-point in these bloods and correlate with tumour burden. To permit quantitative analysis of cfDNA, a qPCR assay targeting 113 bp of *Gapdh* was designed. The assay had a minimum linear dynamic range of 310 – 20,000 pg/reaction, an efficiency of 96% and $R^2$ of 0.998, as determined by a standard curve of serially-diluted genomic DNA (Fig 4.7).
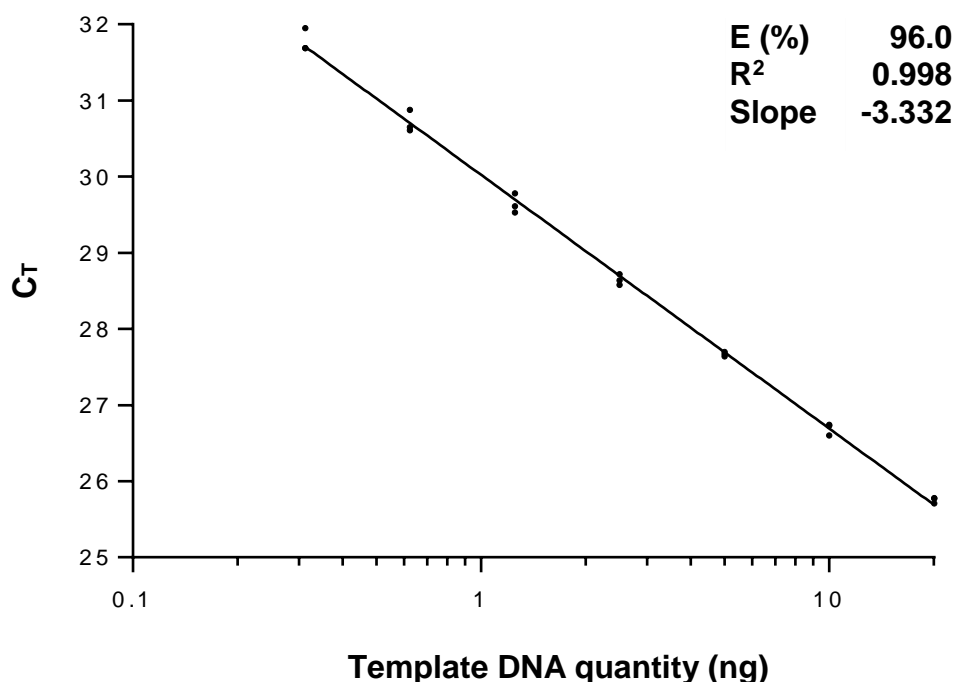
**Figure 4.7 – Standard curve for total quantification of cfDNA by targeting 113 bp of** *Gapdh*.
A TaqMan® qPCR assay targeting a 113 bp region of *Gapdh* was validated through a 7-point, 1:2 dilution series of genomic DNA. Data points represent three replicates per dilution. $C_T$, cycle threshold.

Based on the notion that cfDNA levels are generally elevated in human cancer, it was expected that mice with evidence of the recombined *Kras* allele in cfDNA at end point would have elevated total cfDNA levels (Diaz and Bardelli, 2014). cfDNA was quantified by qPCR (Fig 4.7) from cardiac blood along with 7 WT adenoviral Cre-treated mice (Fig 4.8). Tumour-bearing Kras[+/LSL-G12D] mice exhibited significantly greater cfDNA levels (mean, 2018.0 GE/mL) than WT mice (mean, 763.4 GE/mL) ($P$ = 0.006). There was no correlation between presence/absence of the recombined *Kras* allele in cfDNA (Fig 4.8) and level of total cfDNA. To assess for a possible correlation between end-point total cfDNA levels and tumour burden, H/E lung sections were prepared for each mouse and tumour burden was calculated with ImageJ Software. There was no significant correlation between end-point cfDNA levels and lung tumour burden ($R^2$ = 0.05) (Fig 4.9), which may have been due to the limited assessment of tumour burden in a single H/E slide rather than throughout the whole lung.
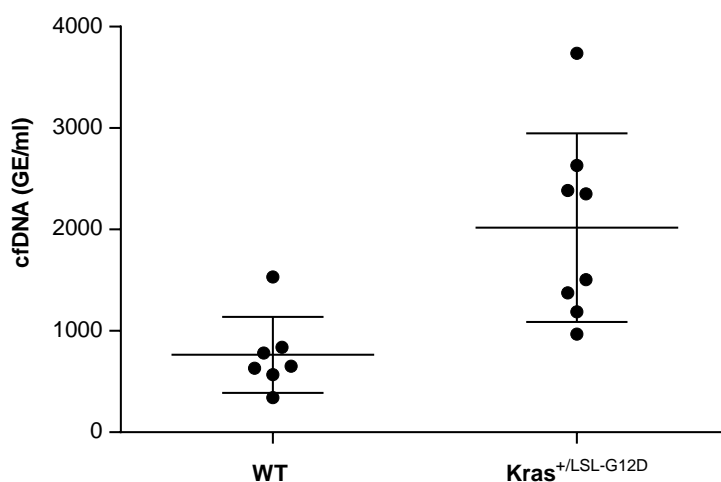
**Figure 4.8 – cfDNA levels at end-point are significantly elevated in tumour-bearing Kras[+/LSL-G12D] mice relative to WT mice.** Using the *Gapdh* qPCR assay in Figure 4.7, cfDNA was quantified from plasma isolated from cardiac blood in 8 Kras[+/LSL-G12D] mice and 7 WT adenoviral Cre-treated control mice. GE, genome equivalent (3.3 pg)
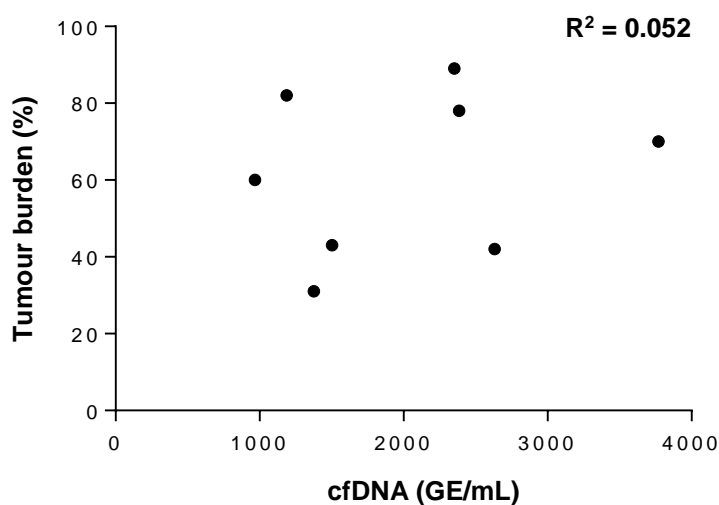


**Figure 4.9 – cfDNA levels were not correlated with tumour burden in Kras[+/LSL-G12D] mice 12 weeks post-infection with 1 x 10[8] PFU adenoviral Cre.** cfDNA was quantified from end-point blood using the *Gapdh* assay in Figure 4.7. Tumour burden was calculated with ImageJ by taking the ratio of tumour to normal area on a single H/E slide. GE, genome equivalent (3.3 pg).

In summary, analysis of cfDNA in Kras[+/LSL-G12D] mice (Cohort 1) after 12 weeks of tumour progression showed evidence of the recombined *Kras* allele in 3/7 mice. Moreover, tumour-bearing mice had significantly higher levels of total cfDNA at end-point compared to age-matched WT control mice. These findings are novel and encouraging, considering the Kras[+/LSL-G12D] mice used in Cohort 1 did not bear malignant tumours at end-point. Based on the frequency of *BRAF*

and *KRAS* mutations in human NSCLC, it was decided that the Kras$^{+/LSL-G12D}$ model was more clinically relevant than the Braf$^{+/LSL-V600E}$ model, and therefore all subsequent mouse cfDNA work was carried out in the Kras$^{+/LSL-G12D}$ model.

### 4.3.2    Study design

The proof-of-concept work discussed in Chapter 4.3.1 was carried out to determine whether *Braf*$^{Lox-V600E}$ and *Kras*$^{Lox-G12D}$ could be detected in the cfDNA of Braf$^{+/LSL-V600E}$ and Kras$^{+/LSL-G12D}$ mice, respectively, at end-point. After confirming a tumour-derived component of cfDNA in these mice, a series of further studies were conducted in which total cfDNA levels and levels of *Kras*$^{Lox-G12D}$ and *Kras*$^{LSL-G12D}$ were quantified over time until end-point in Kras$^{+/LSL-G12D}$ mice.

Mice were scanned at baseline and at 5-week intervals until end-point with a Quantum FX μCT Imaging System. Respiratory gating software was used to correct for breathing-related artefacts. Small (40 μL) blood samples were taken from the saphenous vein at each time point, and an end-point blood sample (200 – 1000 μL) was taken from the heart under general anaesthetic. Since the tumour burden of mice in the proof-of-concept study was high and therefore not representative of a typical tumour burden in a human lung, the titre of adenoviral Cre administered to subsequent mice was halved to 5 x 10$^7$ PFU. Mice were aged to 18 weeks or were sacrificed when pre-defined symptom criteria were met (Chapter 2.7.4).

To avoid blood clotting, which would complicate the separation of plasma from the cellular fraction of centrifuged blood, saphenous blood was aspirated from mice with a pipette tip coated with a solution of 4 mM EDTA in PBS. To facilitate plasma isolation from the 40 μL blood sample, the blood was mixed with 200 μL of 4 mM EDTA solution before centrifugation. End-point cardiac blood was drawn into an EDTA S-Monovette using the aspiration method, and plasma isolated with the standard protocol.

A schematic overview of the studies is shown in Figure 4.10.
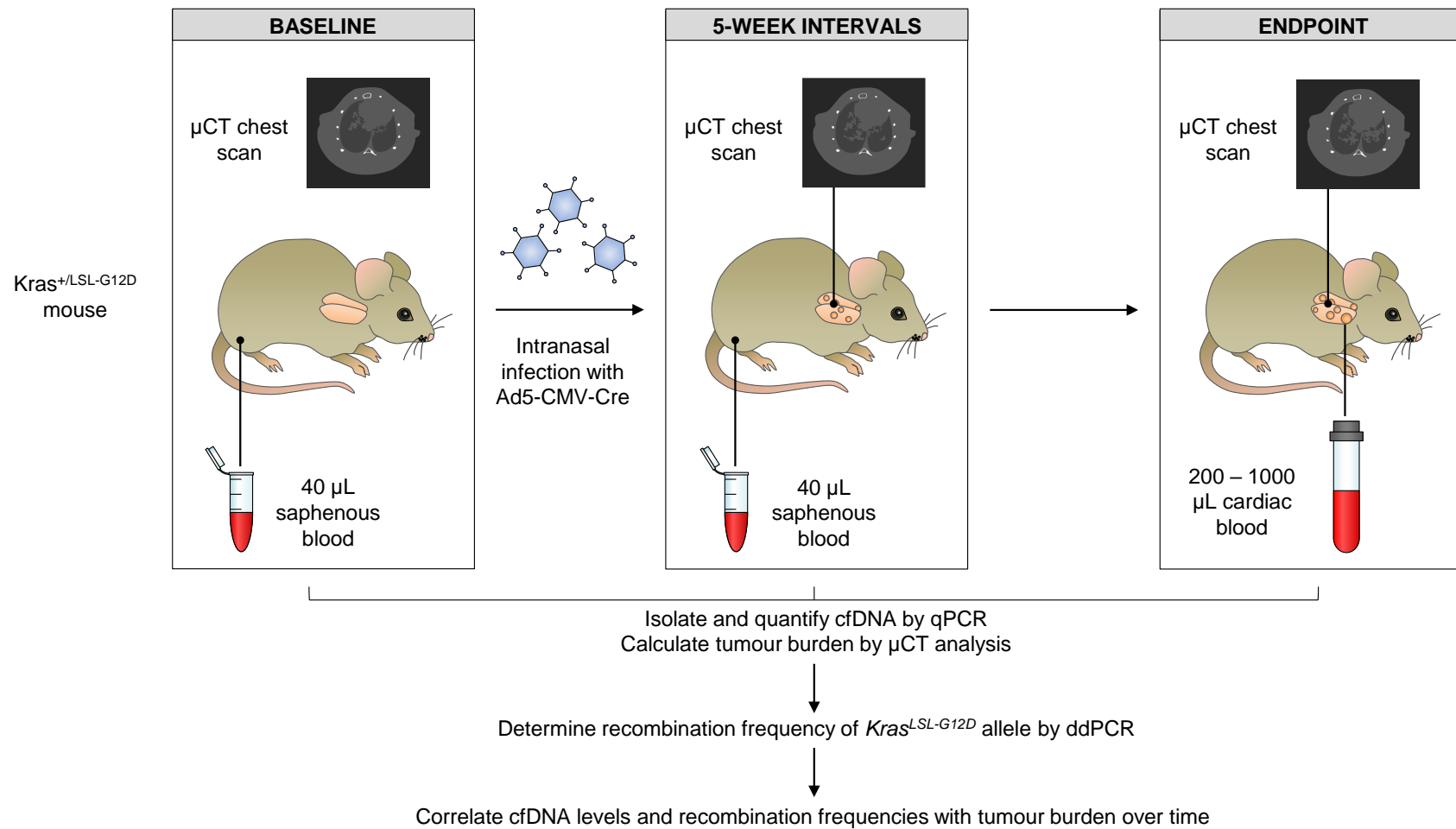
**Figure 4.10 – Schematic overview of cfDNA study in Kras**[+/LSL-G12D] **mice.**

### 4.3.3 Non-invasive analysis of lung tumour burden

Histological analysis of lung tissue in the Kras[+/LSL-G12D] mouse for comparison with cfDNA requires sequential harvesting of large animal cohorts at defined time points. In addition to ethical concerns, this method cannot account for the asynchronous nature of tumour progression between mice, and assumes that tumours always arise from the same lung regions and exhibit the same biological behaviour, thus yielding incoherent results. Micro computed tomography (μCT) imaging has been reported as an alternative tool for the non-invasive monitoring of tumour burden in conditional mouse models, including the Kras[+/LSL-G12D] model (Haines *et al.*, 2009; Lalwani *et al.*, 2013).

Work carried out by a graduate colleague, Mr Callum Rakhit (MRC Toxicology Unit), showed that in the Kras[+/LSL-G12D] mouse lung, volumes of the largest tumours can be estimated by histological analysis of serially-sectioned lung tissue and are significantly correlated with tumour volumes calculated by μCT (n = 4, $R^2$ = 0.96). These data led to the design of further studies in which mice were imaged at 5-week intervals from baseline until end-point. However, identification and quantification of the majority of tumours posed significant challenges. The widespread alveolar hyperplasia that is characteristic of the Kras[+/LSL-G12D] mouse lung (Fig 4.11) produced a fogged appearance upon μCT analysis in the later stages of disease, and structures such as vasculature and bronchioles are of a similar density to tumour tissue (Haines *et al.*, 2009; Lalwani *et al.*, 2013).
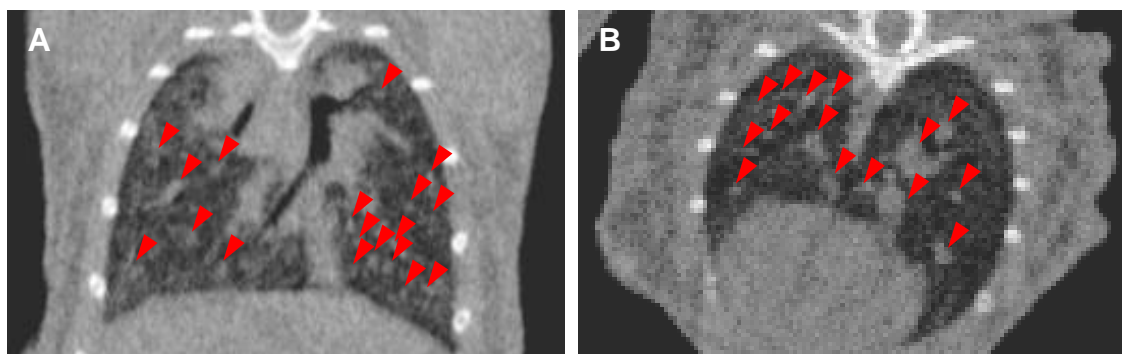


**Figure 4.11 – Representative appearance of Kras[+/LSL-G12D] mouse lungs via μCT.** (A) coronal and (B) axial lung sections at 20 weeks post-adenoviral Cre infection. Red triangles point to non-tumour structures that confound analysis of individual tumours; no tumours can be visualised in these virtual sections.

To overcome issues with identification of individual tumours, measurement of total airspace ('functional lung volume') was investigated at end-point in WT (n = 11) and mutant (n = 16) mice. Measurements were more dispersed for mutant mice relative to WT mice, but there was no significant difference in functional lung volume between groups ($P$ = 0.92; Fig 4.12 A). Analysis of functional lung volume at 5-week intervals in both WT and mutant revealed an increase over time (Fig 4.12 B, C), suggestive of normal age-related growth.

Since measurement of functional lung volume did not show significant differences between WT and mutant mice, both during tumour progression and at end-point, an alternative quantification approach was sought. In order to provide a measure of tumour burden, a quantification approach was taken from Haines *et al*, in which *functional* lung volume was subtracted from *total* lung volume to give a measurement of *tumour and vasculature* (T+V) (Haines *et al.*, 2009). T+V measurement was considered a good estimate of tumour burden and was expressed as a percentage of total lung volume. When measured at end-point (20 weeks post-adenoviral Cre infection), T+V was significantly higher in mutant mice (mean, 22.5%) relative to WT mice (mean, 15.3%) ($P$ = 0.004, Fig 4.13 A). Analysis of %T+V over time showed an increase in mutant mice, but not WT mice, confirming the observed difference in %T+V between groups was indeed tumour-specific (Fig 4.13 B, C). This quantification method therefore provides a non-invasive assessment of disease burden over time, which may be correlated with levels total cfDNA and levels of the recombined (*Kras*[Los-G12D]) allele.
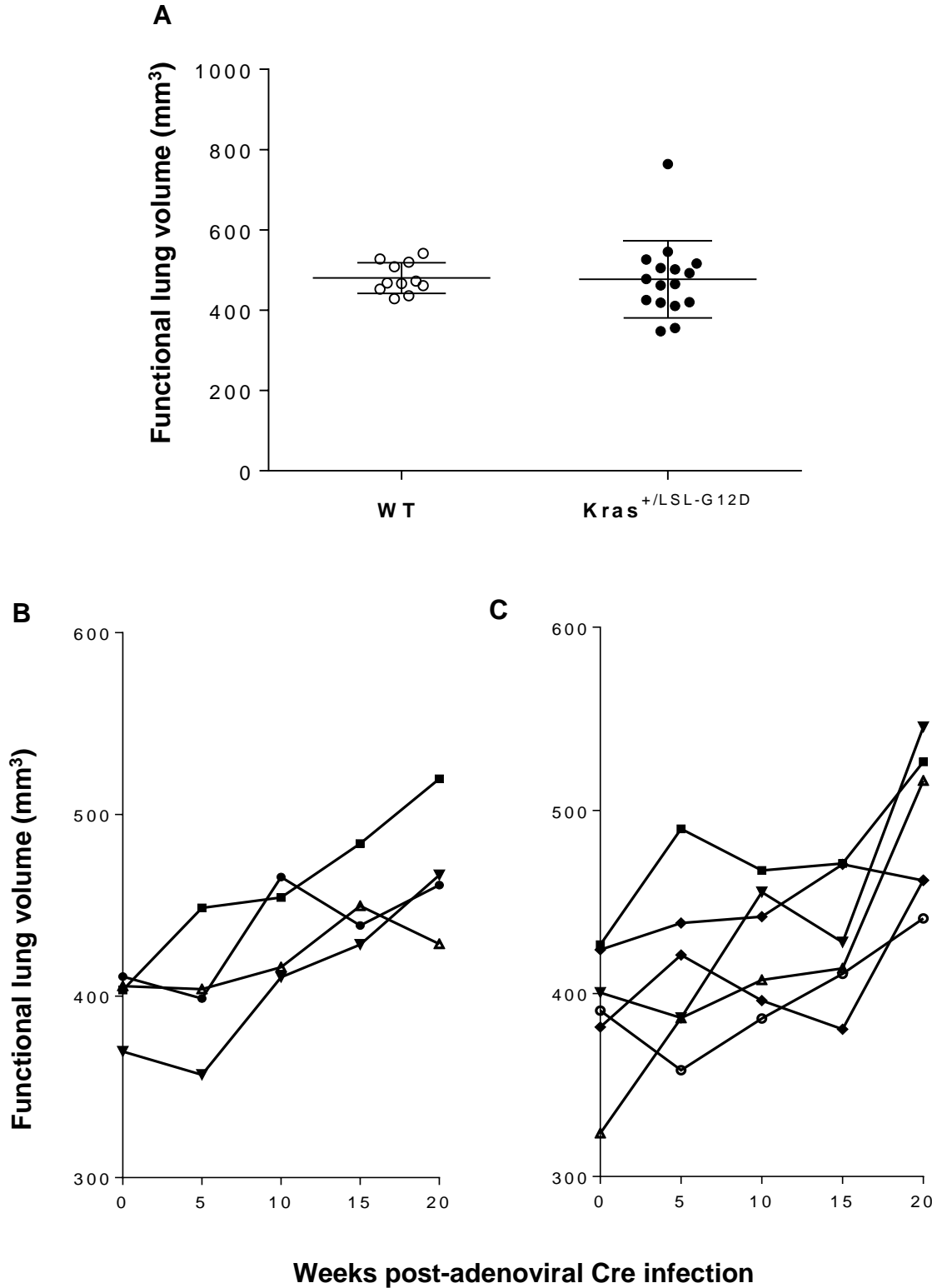
**Figure 4.12 – Functional lung volume analysis in WT and Kras[+/LSL-G12D] mouse lungs at end-point (A) and longitudinally (B, C) via CT imaging.** (A) Functional lung volume was calculated via CT imaging at end-point (20 weeks post-adenoviral Cre infection) in WT and Kras[+/LSL-G12D] mice. (B) Functional lung volume at 5-week intervals in WT mice. (C) Functional lung volume at 5-week intervals in Kras[+/LSL-G12D] mice.
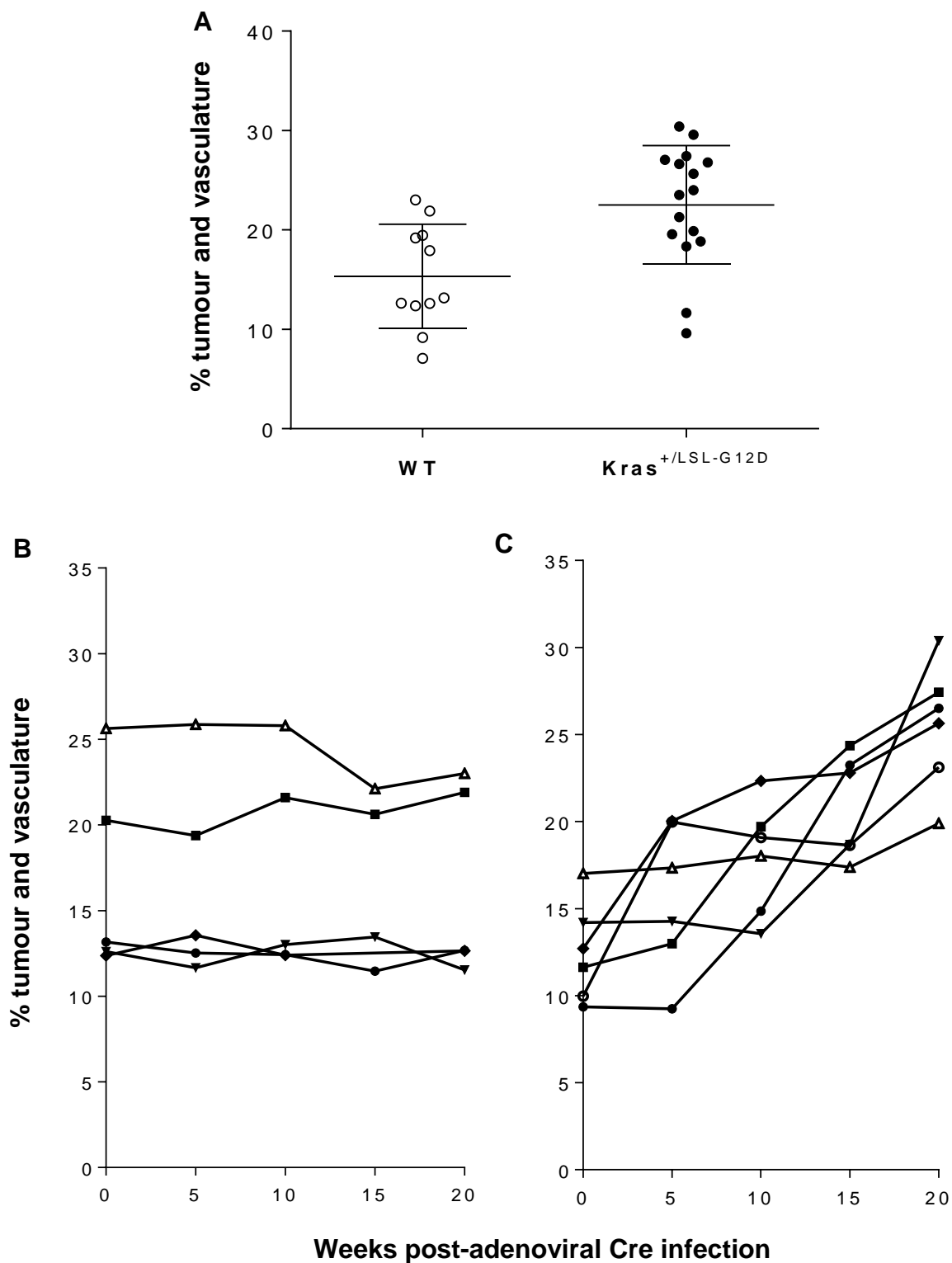
**Figure 4.13 – '% tumour and vasculature' analysis in WT and Kras[+/LSL-G12D] mouse lungs at end-point (A) and longitudinally (B, C) via CT imaging.** (A) '% tumour and vasculature' was calculated via CT imaging at end-point (20 weeks post-adenoviral Cre infection) in WT and Kras[+/LSL-G12D] mice. (B) '% tumour and vasculature' at 5-week intervals in WT mice. (C) Functional lung volume at 5-week intervals in Kras[+/LSL-G12D] mice.

### 4.3.4    Longitudinal monitoring of total cfDNA levels

To correlate tumour burden with the cfDNA profile over time, 40 µL blood samples were taken from the saphenous vein at 5-week intervals from baseline to 15 weeks. Around 200 µL cardiac blood samples were obtained at 20 weeks (end-point). Plasma was isolated and stored at -80˚C and cfDNA was extracted from all samples at the end of the study to avoid degradation-related bias in qPCR analysis. cfDNA was quantified with the 70 bp *Gapdh* pseudogene qPCR assay (Fig 3.4).

There was no trend in cfDNA levels over time in both WT and mutant mice (Table 4.1) and no significant difference in cfDNA levels at any time point between WT and mutant mice ($P > 0.05$). Levels were typically between 0 and 20 pg/µL at all time points (0, 5, 10, 15 and 20 weeks), which is in the same order of magnitude as those from end-point (12 weeks) bloods in WT and mutant mice from the proof-of-concept study (Fig 4.8) (combined range, 2.3 – 25 pg/µL). Some samples from both WT and mutant mice had cfDNA levels greater than 100 pg/µL (highlighted in green in Table 4.1). The plasma samples giving unexpectedly high DNA yields generally had a red tinted appearance, indicating haemolysis and possible lysis of leukocytes; therefore, the most likely explanation for cfDNA levels greater than 100 pg/µL is genomic DNA contamination, which is not tumour-specific. Coagulation was observed in five blood samples (3 longitudinal and 2 end-point), indicated by the dashed boxes in Table 4.1. cfDNA was not quantified in these samples, due to difficult in isolating the aqueous phase from blood, and also due to possible release of genomic DNA from leukocytes.

Given the small volumes of blood collected, the lack of trend in cfDNA levels over time could be related to inefficient DNA extraction, since the Qiagen kits are optimised for use with 200 µL plasma. Also, the interface between the aqueous and cellular fractions of centrifuged blood was minimal, possibly resulting in aspiration of the buffy coat layer in some samples. Based on prior experience in collecting blood from the saphenous vein of C57BL/6J mice, samples greater than 40 µL cannot be reproducibly collected for longitudinal analysis, and so the issue of limited sample volume could not be obviated.

| ID | Genotype | Total cfDNA level (pg/µL blood) | | | | |
|---|---|---|---|---|---|---|
| | | Week 0 | Week 5 | Week 10 | Week 15 | Week 20 |
| 1 | +/+ | 10.1 | 15.8 | / | 10.0 | 25.6 |
| 2 | +/+ | 5.1 | 15.3 | 18.2 | 10.3 | 10.0 |
| 3 | +/+ | 11.4 | 124.4 | 19.7 | 16.3 | 15.7 |
| 4 | +/+ | 6.3 | 7.8 | 35.8 | 3.3 | 17.1 |
| 5 | +/+ | 99.1 | 5.0 | 148.4 | 0.4 | 0.8 |
| 6 | +/+ | 5.0 | 14.7 | 3.5 | 0.2 | 8.9 |
| 7 | +/+ | 11.2 | 3.2 | 7.7 | 0.5 | 17.7 |
| 8 | +/+ | 2.1 | 7.4 | 11.0 | 6.3 | 11.5 |
| 9 | +/+ | 0.9 | 2.6 | 6.3 | 4.6 | 150.4 |
| 10 | +/+ | 232.0 | 7.1 | 8.2 | 12.2 | 0.9 |
| 11 | +/LSL-G12D | 22.2 | 17.5 | 11.7 | 7.2 | 5.3 |
| 12 | +/LSL-G12D | 49.4 | 21.8 | 18.6 | 11.8 | 15.7 |
| 13 | +/LSL-G12D | 19.1 | 8.5 | 8.4 | 0.1 | 0.8 |
| 14 | +/LSL-G12D | 55.4 | 10.1 | 5.8 | 0.9 | 17.7 |
| 15 | +/LSL-G12D | 117.9 | / | / | 29.3 | 150.4 |
| 16 | +/LSL-G12D | 23.2 | 4.2 | 6.6 | 1.3 | 0.9 |
| 17 | +/LSL-G12D | 9.6 | 9.2 | 5.5 | 5.6 | 11.5 |
| 18 | +/LSL-G12D | 16.6 | 24.4 | 11.1 | 0.4 | 8.1 |
| 19 | +/LSL-G12D | 3.8 | 165.5 | 5.5 | 0.9 | / |
| 20 | +/LSL-G12D | 2.3 | 3.1 | 1.9 | 2.2 | 1.3 |
| 21 | +/LSL-G12D | 5.6 | 5.9 | 4.2 | / | 4.8 |

**Table 4.1 – Longitudinal analysis of total cfDNA levels in Kras**[+/+] **and Kras**[+/LSL-G12D] **mice by qPCR**. Total cfDNA levels were measured from 40 µL saphenous blood samples at weeks 0, 5, 10 and 15, and from 200 µL cardiac blood samples at week 20 (end-point), using the 70 bp *Gapdh* pseudogene qPCR assay targeting 44 genomic loci (Fig 3.4). Green shaded boxes highlight samples with cfDNA concentration > 100 pg/µL, and dashed boxes indicate samples that could not be quantified due to visible blood coagulation.

## 4.3.5      Quantitative detection of recombined *Kras* (*Kras*<sup>Lox-G12D</sup>)

In the proof-of-concept study, detection of *Kras*<sup>+/Lox-G12D</sup> in cfDNA of mice at end-point clearly demonstrated a tumour-derived component (Fig 4.6). To accurately evaluate changes in the cfDNA profile over time, qPCR and ddPCR assays for detection of the unrecombined (Fig 3.8) and recombined (Fig 3.26) *Kras* alleles were developed, in addition to qPCR assays for the quantification of total cfDNA (Figs 3.4 and 4.7).

Since the assays for the unrecombined and recombined alleles targeted overlapping genomic regions, these assays could not be used in duplex. Considering the limited quantity of cfDNA obtainable from end-point blood samples, singleplex PCR was not feasible. An alternative approach was taken in which the single locus *Gapdh* assay (Fig 4.7) was duplexed with the assay for the recombined allele by ddPCR (Fig 3.26). Since the Kras<sup>+/Lox-G12D</sup> model is heterozygous for the transgene, and the transgene may be either recombined or unrecombined, the copy number of *Gapdh* was halved to estimate the total copy number of the transgenic allele (recombined and unrecombined). The recombined signal was expressed as a percentage of this value to obtain a recombination frequency:

$$\text{recombination frequency } (\%) = \left( \frac{Kras^{Lox-G12D}}{\frac{Gapdh}{2}} \right) \times 100$$

cfDNA was isolated from end-point plasma of 9 Kras<sup>+/+</sup> mice and 15 Kras<sup>+/LSL-G12D</sup> mice (20 weeks post-infection with adenoviral Cre recombinase) and subjected to ddPCR targeting *Kras*<sup>+/Lox-G12D</sup> and *Gapdh* (Table 4.2). Unexpectedly, 5/9 Kras<sup>+/+</sup> mice showed evidence of the recombined allele in plasma, and 9/15 Kras<sup>+/LSL-G12D</sup> mice showed greater levels of *Kras*<sup>+/Lox-G12D</sup> than *Gapdh*, which is biologically impossible. Taken together, this suggests contamination of cfDNA samples with the *Kras*<sup>Lox-G12D</sup> allele. The origin of this is not known, but likely arose through the cloning study described in Chapter 3.3.7 or possibly through general PCR product contamination. Therefore, the recombination frequency in cfDNA could not be correlated with tumour burden. Given the results of the proof-

of-concept work in Chapter 4.3.1.2, demonstrating tumour-derived cfDNA in *Kras*[+/LSL-G12D] mice at 12 weeks post-administration of adenoviral Cre recombinase, it might be expected that cfDNA samples from subsequent mice would contain *Kras*[+/Lox-G12D] at end-point.

| Mouse Genotype | Positive events (ddPCR) | | Recombination frequency (%) |
|---|---|---|---|
| | *Kras*[Lox-G12D] | *Gapdh* | |
| Kras[+/+] | 0 | 0 | - |
| Kras[+/+] | 3 | 3 | - |
| Kras[+/+] | 29 | 2 | - |
| Kras[+/+] | 11 | 5 | - |
| Kras[+/+] | 35 | 2 | - |
| Kras[+/+] | 4 | 14 | 57.1 |
| Kras[+/+] | 0 | 18 | 0.0 |
| Kras[+/+] | 0 | 0 | - |
| Kras[+/+] | 0 | 0 | - |
| Kras[+/Lox-G12D] | 7 | 6 | - |
| Kras[+/Lox-G12D] | 1 | 0 | - |
| Kras[+/Lox-G12D] | 14 | 2 | - |
| Kras[+/Lox-G12D] | 15 | 2 | - |
| Kras[+/Lox-G12D] | 7 | 6 | - |
| Kras[+/Lox-G12D] | 4 | 3 | - |
| Kras[+/Lox-G12D] | 5 | 63 | 15.9 |
| Kras[+/Lox-G12D] | 3 | 3 | - |
| Kras[+/Lox-G12D] | 0 | 2 | 0.0 |
| Kras[+/Lox-G12D] | 64 | 17 | - |
| Kras[+/Lox-G12D] | 3 | 2 | - |
| Kras[+/Lox-G12D] | 6 | 127 | 9.4 |
| Kras[+/Lox-G12D] | 0 | 0 | - |
| Kras[+/Lox-G12D]* | **283** | **538** | **105.2** |
| Kras[+/LSL-G12D]* | **1** | **505** | **0.4** |

**Table 4.2 – Detection of *Kras*[+/Lox-G12D] allele in cfDNA of mice at end-point by ddPCR**. Blood was obtained by cardiac puncture from 15 Kras[+/LSL-G12D] mice and 9 Kras[+/+] mice at 20 weeks post-infection with $5 \times 10^7$ PFU adenoviral Cre recombinase. Plasma was isolated, cfDNA extracted, and ddPCR assays targeting *Kras*[Lox-G12D] and *Gapdh* run on cfDNA. *MEF cell line.

## 4.4      Discussion

Two-thirds of NSCLC patients present with metastatic spread that is refractory to curative treatment (Molina *et al.*, 2008). Despite intense efforts to evaluate the utility of cfDNA in NSCLC, late clinical presentation of the disease hampers the ability to study early-stage disease in humans. Hence, the earliest stage of the disease at which cfDNA has diagnostic potential is currently unknown. Several studies have demonstrated a tumour-derived component of cfDNA in xenografted mice bearing small tumours. However, xenograft mice are not suitable for the study of cfDNA in early-stage disease because the xenografted cells are transformed prior to implantation; thus, tumour development is not strictly *de novo*. Also, xenografted tumours lack an immune component, the importance of which has been highlighted in recent years. Some genetically engineered mouse (GEM) models are effective at recapitulating the early stages of human tumourigenesis, and may be suitable for analysis of cfDNA in early-stage disease (Chapter 1.9).

In this chapter, the Kras^+/LSL-G12D^ mouse was used to model lung preneoplasia and permit analysis of cfDNA from the earliest stages of development. The overall aim was to determine the earliest stage of development at which tumour-derived DNA was detectable in cfDNA, to inform the diagnostic potential of cfDNA in the cognate human disease. This work follows the proof-of-concept study in which the recombined *Kras* allele (*Kras*^Lox-G12D^) was identified in the blood of Kras^+/LSL-G12D^ mice with lung preneoplasia 12 weeks post-infection with adenoviral Cre, using end-point PCR (Chapter 4.3.1.2). In addition, total cfDNA levels at end-point were significantly higher in tumour-bearing mice than in WT control mice.

### 4.4.1      Limitations of adenoviral Cre recombinase administration

A recent study by Sutherland *et al* highlighted the importance of restricting Cre expression to specific cell types in the lung; using several cell type-specific adenoviral vectors, the authors were able to induce the development of adenocarcinomas from multiple cells of origin in the Kras^+/LSL-G12D^ mouse.

Moreover, the cells of origin determined the histopathological features of the resulting tumours, suggesting lung adenocarcinoma may originate from multiple cell types (Sutherland *et al.*, 2014). However, in the current study, adenoviral Cre recombinase (Ad5-CMV-Cre) was expressed under the control of the ubiquitous CMV promoter, which enabled Cre-mediated recombination in a cell type-independent manner. Due to the route of viral administration, and the indiscriminate expression of Cre, it is possible that some mice could have developed undetected nasopharyngeal neoplasia that may have contributed to the circulating *Kras*$^{Lox\text{-}G12D}$ allele fraction. Unpublished data from Prof Pritchard's lab has also shown that ubiquitous Cre is capable of recombining myeloid-lineage cells of the lung, raising further concern over the origin(s) of the circulating *Kras*$^{Lox\text{-}G12D}$ allele fraction in the Kras$^{+/LSL\text{-}G12D}$ model. However, the possibility of Cre-mediated recombination in blood cells was excluded in Chapter 4.3.1.2. A more reliable way of activating the conditional allele in lung cells would be through use of a promoter-specific adenoviral Cre, e.g. SPC-adenoviral Cre.

### 4.4.2      Longitudinal monitoring of tumour burden

The Kras$^{+/LSL\text{-}G12D}$ mouse is characterised histologically by widespread alveolar hyperplasia and multiple adenomas, predominantly with bronchiolar involvement. Most adenomas do not progress to malignancy over the time period that mice develop symptoms of respiratory stress. Therefore, the majority of mice in this study did not develop overt adenocarcinomas. Quantitative histological analysis of Kras$^{+/LSL\text{-}G12D}$ lung tissue at end-point was challenging due to widespread alveolar hyperplasia and diffuse tumour borders; ImageJ was used to calculate total tumour area, but likely provided an underestimate due to the presence of diffuse hyperplasia. Moreover, histological analysis of tumour burden in the Kras$^{+/LSL\text{-}G12D}$ mouse based on single slides is subject to significant sampling bias because of the bronchiolar origin of tumours.

To permit a longitudinal assessment of tumour burden in the Kras$^{+/LSL\text{-}G12D}$ mouse and reduce animal usage, μCT imaging was evaluated. Data generated by a

colleague, Mr Callum Rakhit (MRC Toxicology Unit, Leicester), shows that in the Kras$^{+/LSL-G12D}$ mouse lung, volumes of the largest tumours can be estimated by histological analysis of serially-sectioned lung tissue and are concordant with tumour volumes calculated by μCT. These data led to the design of further studies in which mice were imaged at 5-week intervals from baseline until end-point.

The μCT imaging data was initially analysed for total lung airspace volume and individual tumour volume. However, accurate quantification of tumour volume was met with several challenges. Firstly, as previously reported, tumour borders in the Kras$^{+/LSL-G12D}$ mouse were often diffuse, and therefore the contrast in density between tumour and normal regions was low (Haines *et al.*, 2009; Lalwani *et al.*, 2013). Secondly, diffuse alveolar hyperplasia throughout the lung masked over some tumours, particularly in later scans and the end-point scan where lungs exhibited a 'fogged' appearance. Thirdly, airway and vascular structures in the lung had a similar density to tumour tissue, which confounded identification of tumour, as did the presence of multiple tumours in close proximity (Haines *et al.*, 2009; Lalwani *et al.*, 2013). These factors were further complicated by bronchiolar origin of most tumours. GEM models of lung cancer in which oncogenes are activated by spontaneous recombination (e.g. the Kras$^{LA1}$ mouse that expresses *Kras*$^{G12D}$) do not develop tumours associated with bronchioles, and may therefore be more amenable to imaging (Badea *et al.*, 2012; Johnson *et al.*, 2001).

Future volumetric analyses of lung tumours could benefit from the use of contrast-enhancing agents. Several studies have demonstrated the ability of injectable iodine-based or nanoparticulate agents to increase contrast between tumour and vascular tissue in the Kras$^{+/LSL-G12D}$;Trp53$^{Fl/Fl}$ mouse, improving tumour margins and permitting more accurate quantification (Ashton *et al.*, 2014; Lalwani *et al.*, 2013; Badea *et al.*, 2012). Data generated by Mr Callum Rakhit shows how opacifying vascular tissue (using eXIA$^{TM}$ 160 contrast agent) improves the identification of tumours in the Kras$^{+/LSL-G12D}$ and Kras$^{+/LSL-G12D}$;Trp53$^{Fl/Fl}$ models.

An alternative approach to improve soft-tissue contrast in the Kras$^{+/LSL-G12D}$ mouse would be through use of bioluminescence. Rodriguez *et al.* developed a lentivirus-based approach to incorporate luminescence into Cre-Lox-regulated

140

mouse models. The lentivirus contains a transgenic construct comprising luciferase and Cre-ER[T] genes, which are coded on opposite strands, as shown in Fig 4.7. Both genes are under the control of a single CAG promoter. The promoter is floxed such that, in the presence of tamoxifen, Cre recombinase inverts the promoter sequence, permitting expression of luciferase. The conditional allele in a cell infected by the lentivirus (or adenovirus) will be Cre-recombined prior to silencing of Cre. This strategy could be readily used with the Kras[+/LSL-G12D] mouse, since it does not require any germline modification (Rodriguez *et al.*, 2014). Tumour-specific luminescence can also be achieved by positioning the reporter gene directly upstream or downstream of an oncogene, or by crossing a luminescent strain with a cancer strain (Ju *et al.*, 2015; Woolfenden *et al.*, 2009).
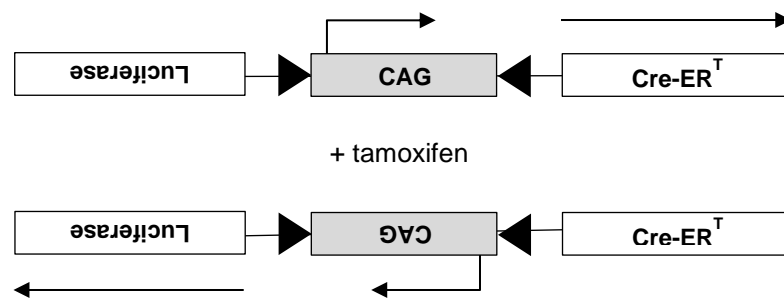


**Figure 4.7 – Luciferase construct for introduction of bioluminescence to a Cre-Lox-regulated mouse model.**

Since contrast between tumour and vascular tissue was low in the Kras[+/LSL-G12D] mouse, an alternative approach was taken from Haines *et al.* in which the volume of total tumour and vasculature was calculated and expressed as a proportion of whole lung volume. Using this method, tumour burden did increase over time in Kras[+/LSL-G12D] mice but not Kras[+/+] mice. In summary, whilst µCT imaging samples the entire lung, it currently fails to provide an accurate assessment of hyperplasia, and individual tumours are difficult to follow longitudinally with both automated and user-guided analysis pipelines. Without the use of contrast enhancement, analysis is limited to measurement of overall tumour burden.

### 4.4.3 Longitudinal sampling and processing of blood

Longitudinal blood samples were taken from the saphenous vein at 5-week intervals from baseline. To reduce variability in the volume of blood obtained between sampling time points, a volume of 40 µL blood was chosen. Blood was aspirated from a hind leg using an EDTA-coated pipette tip and transferred to 180 µL PBS containing 4 mM EDTA. The small sampling volume presented several methodological challenges. Firstly, 40 µL blood produced a very small interface between the plasma and cellular fractions after centrifugation, increasing the chance of genomic DNA contamination from leukocytes during plasma isolation. Secondly, preventing coagulation was difficult because the saphenous blood was transiently exposed to air. Commercial blood collection devices that contain anticoagulant could not be used because they are typically optimised for larger blood volumes, and hence do not efficiently recover 40 µL of blood. Thirdly, commercial kits commonly used to isolate cfDNA are optimised for larger volumes of plasma (around 200 µL).

Total cfDNA levels from longitudinal sampling did not appear to change significantly over time in any mouse. Whilst this could be a true biological observation, the variation could possibly be due to the influence of blood coagulation and leukocyte contamination. Also, blood coagulation occurred within the pipette tip during aspiration on some occasions, which prevented collection of the full 40 µL sample. Thus, collection and processing of small blood volumes for cfDNA analysis remains a critical challenge that was not solved in this study.

### 4.4.4 Correlating cfDNA levels with tumour burden

Total cfDNA level is a potential parameter for the monitoring of tumour burden in the Kras$^{+/LSL\text{-}G12D}$ mouse. qPCR strategies for the assessment of total cfDNA levels in human are widely reported in the literature, and are based on targeting of single-locus genes, e.g. *GAPDH* and *RPPH1*, or repetitive elements, e.g. *ALU* and *LINE-1*. In this chapter, a qPCR assay targeting single-locus *Gapdh* was developed for the quantification of concentrated mouse DNA. This assay had a

minimum linear dynamic range of 30 – 20,000 pg/reaction. Based on preliminary mouse experiments in which total cfDNA isolated from small (approximately 40 µL) volumes of mouse blood was of very low concentration (5 – 10 pg/µL eluate), a more sensitive qPCR assay was developed. This assay targeted 40 *Gapdh* pseudogenes, enabling a greater linear dynamic range of 2 – 5,000 pg/reaction to be achieved.

The lack of correlation between cfDNA levels and tumour burden in Chapter 4.3.1.2 could have been due to tissue sampling bias and also biological characteristics of the lesion types identified. Whereas in solid human cancers there is often a single histologically-confirmed lesion, the Kras[+/LSL-G12D] model develops a range of lesions that progress asynchronously, and the relative contribution of each lesion type to the overall cfDNA profile is currently unknown. The overwhelming majority of lesions identified in the Kras[+/LSL-G12D] mouse are premalignant and it is currently unclear whether premalignant lesions contribute significantly to cfDNA levels in humans (Hashad *et al.*, 2012; Zhong *et al.*, 2007; Kim *et al.*, 2014).

### 4.4.5    Detection of *Kras*[Lox-G12D] in cfDNA at end-point

The PCR assays developed in Chapter 3 for detection of the unrecombined and recombined alleles of *Kras* are both targeted to overlapping regions of the LoxP sequence, which prevents their use in duplex. Instead, relative levels of unrecombined/recombined allele were estimated by duplexing the assay for the recombined allele with the assay for single-locus *Gapdh*. Since the Kras[+/LSL-G12D] mouse is heterozygous, the concentration of *Gapdh* was halved to estimate the total concentration of the transgenic allele (unrecombined and recombined together). Then, the concentration of recombined allele was combined with the *Gapdh* data to produce an overall recombination frequency. This method was validated with genomic DNA isolated from unrecombined and recombined MEF cell lines, producing recombination frequencies close to 0% and 100%, respectively. Unfortunately, analysis of recombination in cfDNA isolated from

Kras+/LSL-G12D and Kras+/+ mice 20 weeks post-delivery of adenoviral Cre recombinase showed evidence of *Kras*Lox-G12D allele contamination. This issue likely arose from use of a pipette contaminated with the 180 bp LoxP-containing sequence that was cloned in Chapter 3.3.7, and therefore recombination frequencies in cfDNA could not be correlated with tumour burden.

## 4.5        Conclusions

The work in this chapter highlighted several methodological challenges relating to analysis of tumour burden in the Kras+/LSL-G12D mouse, and collection and processing of small blood volumes for analysis of cfDNA. As an alternative to analysis of individual lung lesions over time, overall tumour burden was measured using a previously reported method, which showed a clear increase over time in Kras+/LSL-G12D mice, but not Kras+/+ mice. Measurement of total cfDNA levels over time yielded variable results, and some plasma samples showed obvious haemolysis (red tint). Despite removing outlier samples, there were no trends in cfDNA levels over time in Kras+/LSL-G12D mice or Kras+/+ mice. Unfortunately, tumour burden could not be correlated with levels of recombined allele in cfDNA, because cfDNA extracts were contaminated with a plasmid from the molecular cloning work presented in Chapter 3. Further work should seek to overcome the methodological issues of tumour imaging and blood collection in order to re-address the aims of this chapter.

# Chapter 5

# Analysis of somatic copy number alteration (SCNA) in cfDNA of patients with non-small cell lung cancer

## 5.1        Introduction

Copy number variation (CNV) is a form of natural polymorphic variation in which the copy number status of a particular genomic locus deviates from that observed in a reference genome. CNVs are inherited in the germline, but may also be acquired in somatic cells in cancer as a result of structural deletions, duplications, insertions, inversions and translocations, where somatic copy number alteration (SCNA) is the appropriate term (Hastings *et al.*, 2009). Therefore, SCNA is used throughout this thesis to describe somatic alterations in cancer.

Genome-wide SCNA in cancer can be evaluated with next-generation sequencing and hybridisation-based methods, such as comparative genomic hybridisation (CGH) and SNP microarrays (Shaw *et al.*, 2012; Iranmanesh and Guo, 2014). Discrete genomic loci can be interrogated for SCNA with fluorescence *in situ* hybridisation (FISH), qPCR and dPCR (Vanhecke *et al.*, 2013; Page *et al.*, 2011; Gevensleben *et al.*, 2013). Using qPCR, Page *et al.* detected amplification of *ERBB2*/*HER2* in cfDNA of patients with primary and metastatic breast cancer, suggesting the potential of *ERBB2*) amplification as a circulating biomarker (Page *et al.*, 2011). Gevensleben *et al.* subsequently confirmed this by dPCR with matched tissue from patients with *HER2*-amplified tumours (Gevensleben *et al.*, 2013).

SCNA in specific driver genes characterise the genomic landscape of NSCLC, and amplifications of *EGFR* and *MET* have therapeutic and prognostic significance (Cappuzzo *et al.*, 2005; Hirsch *et al.*, 2005). *EGFR* is mutated in around 15% of NSCLCs in Caucasians and around 35% in East Asians (Jorge *et al.*, 2014). In turn, around half of *EGFR*-mutated tumours also show increased *EGFR* copy number, and *EGFR*-amplified adenocarcinomas are associated with significantly poorer prognosis (Beau-Faller *et al.*, 2008). However, multiple clinical studies have demonstrated that high *EGFR* copy number (as determined by FISH) confers sensitivity to EGFR TKIs and is predictive of improved overall survival (Cappuzzo *et al.*, 2005; Hirsch *et al.*, 2005; Zhu *et al.*, 2008; Hirsch *et al.*, 2006). Amplification of *MET* is recognised as a mechanism of resistance to EGFR TKIs, accounting for up to 20% of resistant tumours, and is observed at low

frequency (approximately 6%) in TKI-naïve tumours. In NSCLC, *MET* amplification often co-exists with *EGFR* amplification and is associated with advanced tumour stages and poor prognosis following surgical resection (Go *et al.*, 2010; Cappuzzo *et al.*, 2009b; Okuda *et al.*, 2008).

Based on the clinical importance of SNCAs in *EGFR* and *MET* NSCLC, it is attractive to speculate that amplification of these genes may be detectable in the cfDNA of patients, and may have utility as circulating biomarkers for therapeutic guidance and disease prognostication. In addition, detection of other amplified genes in cfDNA may aid in the molecular and histological stratification of tumours.

Of particular interest are gene amplifications in preneoplastic squamous lesions of the lung (Chapter 1.6.3), since detection of these events in cfDNA may have relevance to early diagnosis. *PIK3CA* and *SOX2* are both commonly amplified in such lesions, as a gain of chromosome 3q (Massion *et al.*, 2004; Schneider *et al.*, 2013). Moreover, amplification and overexpression of *SOX2* are associated with a more favourable prognosis in SqCC (Wilbertz *et al.*, 2011).

| Gene | Study *n* | | Tumour stage | Frequency of amp (%) | | | Assay | Threshold | Reference |
|------|-----|------|------|-----|------|---------|-------|-----------|-----------|
| | AC | SqCC | | AC | SqCC | Overall | | | |
| *EGFR* | 94 | 89 | I-III | NS | NS | 9 | FISH | GCR ≥ 2.1 | Hirsch *et al.*, 2003 |
| | NS | NS | NS | NS | NS | 9 | FISH | GCR > 2.0 | Cappuzzo *et al.*, 2005 |
| | 128 | 67 | T1-4, N0-2 | 8 | 15 | 10 | FISH | GCR ≥ 2.0 | Dacic *et al.*, 2006 |
| | 49 | 57 | I-IV | 33 | 32 | 32 | qPCR | > M ± 2SD | Beau-Faller *et al.*, 2008 |
| | 99 | 0 | I-IV | 9 | 0 | 9 | FISH | GCR ≥ 2.0 | Sholl *et al.*, 2009 |
| | 241 | 137 | I-IV | NS | NS | 10 | FISH | CN < 5.0 | Cappuzzo *et al.*, 2009b |
| | 133 | 0 | I-IV | 8 | 0 | 8 | FISH | GCR ≥ 2.0 | Liang *et al.*, 2010 |
| *FGFR1* | 77 | 155 | NS | 0 | 10 | 10 | FISH | GCR ≥ 2.0 | Weiss *et al.*, 2010 |
| | 555 | 46 | I-IV | 3 | 21 | 26 | SNPa | CN ≥ 3.25 | Dutt *et al.*, 2011 |
| | 0 | 226 | I-IV | 0 | 16 | 16 | FISH | GCR ≥ 2.2 | Heist *et al.*, 2012 |
| | 0 | 262 | I-III | 0 | 13 | 13 | FISH | GCR > 2.0 | Kim *et al.*, 2013a |
| *HER2* | 21 | 12 | NS | 5 | 0 | 3 | SB | None | Shiraishi *et al.*, 1989 |
| | 24 | 22 | I-IIIa | 8 | 0 | 4 | FISH | GCR > 2.0 | Hirsch *et al.*, 2002 |
| | 143 | 80 | IIIB-IV | 4 | 1 | 3 | FISH | GCR > 3.0 | Heinmoller *et al.*, 2003 |
| | NS | NS | I-III | NS | NS | 22* | FISH | GCR ≥ 3.0 | Pellegrini *et al.*, 2003 |
| | 56 | 47 | III | 5 | 4 | 5 | FISH | GCR > 2.2 | Junker *et al.*, 2005 |
| | NS | NS | NS | NS | NS | 10 | FISH | GCR > 2.0 | Cappuzzo *et al.*, 2005 |
| *MET* | 49 | 57 | I-IV | 21 | 20 | 21 | qPCR | > M ± 2SD | Beau-Faller *et al.*, 2008 |
| | 166 | 0 | III-IV | 7 | 0 | 7 | FISH | GCR ≥ 2.9 | Cappuzzo et al., 2009a |
| | 241 | 137 | I-IV | NS | NS | 4 | FISH | CN < 5.0 | Cappuzzo et al., 2009b |
| | 37 | 7 | I-IIIb | 0 | 14 | 2 | qPCR | > M ± 2SD | Chen *et al.*, 2009 |
| | 25** | 2** | IV | 8 | 50 | 11 | qPCR | > M ± 2SD | Chen *et al.*, 2009 |
| | 72 | 97 | I-IV | 0 | 6 | 4 | FISH | GCR ≥ 2.0 | Go *et al.*, 2010 |
| *PIK3CA* | 64 | 28 | I-IV | 2 | 36 | 12 | qPCR | CN > 3.0 | Kawano *et al.*, 2007 |
| | 94 | 35 | I-IV | 7 | 34 | 15 | FISH | GCR ≥ 2.0 | Okudela *et al.*, 2007 |
| | 195 | 139 | I-IV | 6 | 33 | 17 | qPCR | CN > 4.0 | Yamamoto *et al.*, 2008 |
| | 30 | 66 | I-III | 3 | 42 | 31 | qPCR | CN ≥ 4.0 | Ji *et al.*, 2011 |
| | 68 | 64 | I-IV | 6 | 36 | 19 | FISH | GCR > 2.0 | Wang *et al.*, 2014a |
| *SOX2* | 0 | 47 | I-IV | 0 | 23 | 23 | SNPa | CN > 2.85 | Bass *et al.*, 2009 |
| | 17 | 40 | NS | 0 | 20 | 20 | qPCR | CN ≥ 4.0 | Yuan *et al.*, 2010 |
| | 0 | 136 | NS | 0 | 20 | 20 | aCGH | CN ≥ 2.0 | Hussenet *et al.*, 2010b |
| | 208 | 210 | I-IV | 6 | 76 | 41 | FISH | GCR > 2.0 | Wilbertz *et al.*, 2011 |
| | 291 | 49 | I-IV | 6 | 73 | 15 | FISH | GCR > 2.0 | Wilbertz *et al.*, 2011 |
| | 70 | 77 | I-IIIa | 8 | 72 | 39 | FISH | GCR > 2.0 | Brcic *et al.*, 2012 |
| | 0 | 178 | I-IV | 0 | 21 | 21 | SNPa | NS | TCGA R Network 2012 |
| | 0 | 33 | III | 0 | 22 | 67 | FISH | CN ≥ 2.0 | Yoon *et al.*, 2016 |

**Table 5.1 – Oncogenes subject to frequent amplification in NSCLC**. *(7/41). **TKI-resistant. aCGH, array comparative genomic hybridisation; CN, copy number; FISH, fluorescence *in situ* hybridisation; GCR, gene:chromosome ratio; NS, not specified; -, rare/absent; SB, Southern blot; SNPa, single nucleotide polymorphism array; TKI, tyrosine kinase inhibitor.

In this chapter, SCNA in 6 genes (Table 5.1) showing common amplification in NSCLC was investigated in cfDNA from patients with NSCLC in two clinical trials: the ReSoLuCENT study (NIHR Clinical Research Network, UK) and the GALAXY-1 clinical trial (Madrigal Pharmaceuticals, Inc., USA).

ReSoLuCENT (Resource for the Study of Lung Cancer Epidemiology in North Trent) is an ongoing study that aims to gather epidemiological and biological data from patients with early-stage lung cancer and patients with a family history of the disease. This information will be used to study acquired and inherited genetic changes in lung cancer, and to assist in proteomic analyses. The study has obtained tumour material, serum and plasma. In this chapter, SCNA was investigated in cfDNA and matched genomic DNA (from leukocytes) isolated from plasma samples of patients enrolled on the ReSoLuCENT study and compared with mutation data obtained from Ion Torrent™ NGS.

GALAXY-1 was a randomised, phase IIB/III study that investigated whether combining the heat shock protein 90 (hsp90) inhibitor, ganetespib, with docetaxel is more effective than docetaxel alone, as a second-line treatment in advanced (IIIB/IV) NSCLC. The trial was open to patients with evidence of radiological progression following systemic therapy for advanced disease (Ramalingam *et al.*, 2015). Blood was collected from patients at baseline and during the first and second cycles of treatment, and cfDNA was evaluated with the Ion Torrent™ NGS platform. SCNA analysis in this chapter was performed on baseline cfDNA and compared with mutation data obtained through NGS. In this study, SCNA and mutation data were combined and correlated with patient survival outcomes.

## 5.2        Aims and objectives

The aims of this chapter were to:

i)        investigate SCNA of 6 genes in cfDNA of patients with NSCLC from the ReSoLuCENT study and GALAXY-1 clinical trial; and

ii)        determine whether SCNA and/or mutations in cfDNA are correlated with survival of patients with advanced NSCLC in the GALAXY-1 trial.

The objectives were to:

i)        develop qPCR assays for SCNA analysis;

ii)        validate use of preamplification PCR to enable multi-target analysis

iii)        investigate SCNA in 6 genes in cfDNA of patients with locally advanced/advanced NSCLC;

iv)        compare results by qPCR with droplet digital PCR (ddPCR); and

v)        compare SCNA and mutation data obtained through NGS of cfDNA.

## 5.3        Results

### 5.3.1        Design and validation of qPCR assays for SCNA analysis

Based on a review of the literature and publically available data, qPCR assays were designed as described in Chapter 2.8.1.1 to target *EGFR*, *FGFR1*, *MET*, *MYC*, *MYCN*, *PIK3CA* and *SOX2* (details of primer sequences are given in Table 2.5). The performance of each assay was evaluated by standard curve analysis of serially diluted human genomic DNA (hgDNA) over a range of 0.156 ng to 10 ng (Fig 5.1). Each assay showed linearity within this concentration range, and had an amplification efficiency within acceptable limits (95 - 105%) (Table 5.2). Slopes created by standard curve analysis were parallel for all assays, including the reference (*RPPH1*), which permitted use of the reference assay in combination with all target assays. Figure 5.1 shows a representative example of a standard curve targeting *MET*.  $R^2$, efficiency and slope values are detailed for each assay in Table 5.2.
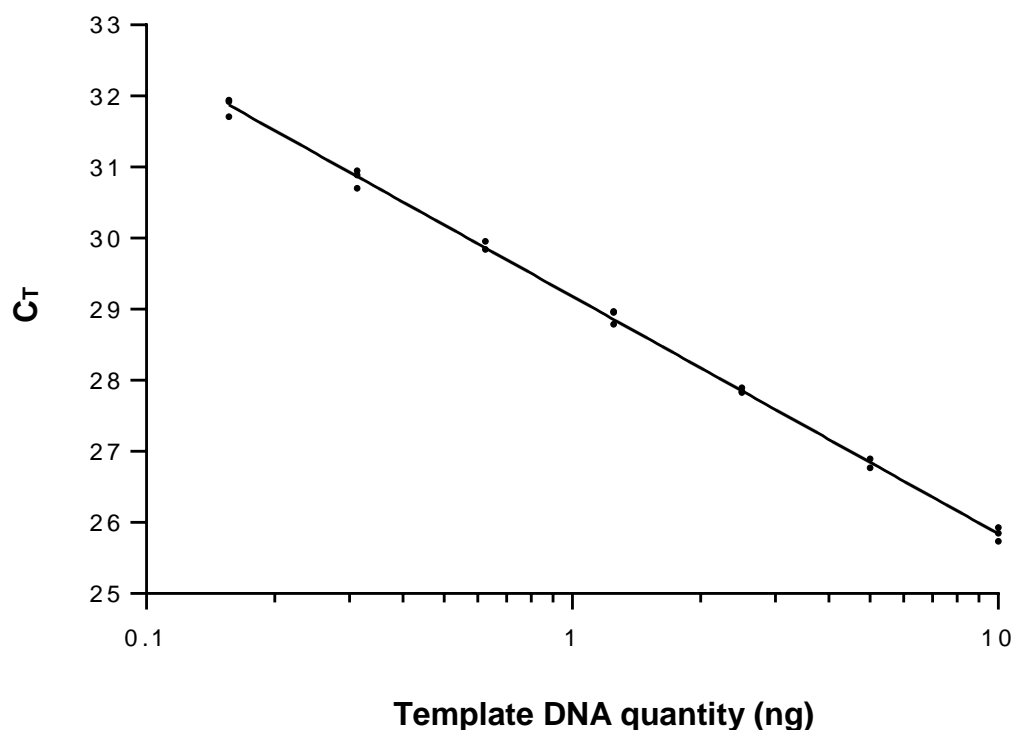


**Figure 5.1 – qPCR SCNA assay evaluation**. Representative standard curve for a qPCR assay targeting *MET*, constructed with a seven-point, 1:2 dilution of hgDNA. Each dilution was performed in triplicate. $C_T$, threshold cycle.

| Assay | R² | E (%) | Slope |
|-------|------|-------|--------|
| *RPPH1* | 0.994 | 98.0 | -3.352 |
| *EGFR* | 0.995 | 99.7 | -3.309 |
| *FGFR1* | 0.993 | 100.5 | -3.221 |
| *HER2* | 0.996 | 99.7 | -3.386 |
| *MET* | 0.998 | 99.3 | -3.349 |
| *PIK3CA* | 0.995 | 102.0 | -3.290 |
| *SOX2* | 0.999 | 98.9 | -3.348 |

**Table 5.2 – Performance parameters of qPCR assays for SCNA analysis**. Parameters were calculated using StepOnePlus™ Software v2.3 (Applied Biosystems). E, efficiency; $R^2$, squared Pearson correlation coefficient.

Having demonstrated that the 7 assays showed similar efficiency, each assay was validated using DNA isolated from cancer cell lines with known gene amplification (Table 5.3). *RPPH1* was chosen as the endogenous reference gene, since it is not associated with CNV or SCNA, and hgDNA as the calibrator, thus enabling relative quantification (RQ).

| Gene | Amplified cell line | Cancer Type | Method | CN | Reference |
|------|---------------------|-------------|--------|------|-----------|
| *EGFR* | HCC827 | Lung AC | FISH | > 20 | Amann *et al.*, 2005 |
| | | | aCGH | 36 | Amann *et al.*, 2005 |
| | PC-9 | Lung AC | FISH | NS | Ohashi *et al.*, 2012 |
| | PC-9/ER | Lung AC | FISH | NS | Ohashi *et al.*, 2012 |
| *FGFR1* | H520 | Lung SqCC | SNPa | 10 | Weiss *et al.*, 2010 |
| | | | qPCR | 5-6 | Dutt *et al.*, 2011 |
| | | | aCGH | 6 | Rooney *et al.*, 2016 |
| | | | | | Rooney *et al.*, 2016 |
| *HER2* | SKBR3 | Breast AC | FISH | 16-22 | Kallioniemi *et al.*, 1992 |
| | | | qPCR | 19 | Whale *et al.*, 2012 |
| | | | dPCR | 14 | Whale *et al.*, 2012 |
| | | | qPCR | 13 | eBioscience, 2016 |
| | | | FISH | 14 | eBioscience, 2016 |
| *MET* | None available | - | - | - | - |
| *PIK3CA* | H520 | Lung SqCC | FISH | 5-10 | Singh *et al.*, 2002 |
| *SOX2* | HCC95 | Lung SqCC | SNPa | 4.7 | Watanabe *et al.*, 2014 |
| | H520 | Lung SqCC | NS | NS | Bass *et al.*, 2009 |

**Table 5.3 – Cell lines with known gene amplifications used for validation of qPCR SCNA assays**. AC, adenocarcinoma; aCGH, array comparative genomic hybridisation; ER, erlotinib-resistant; FISH, fluorescence *in situ* hybridisation; NS, not specified, SNPa, single nucleotide polymorphism array; SqCC, squamous cell carcinoma.
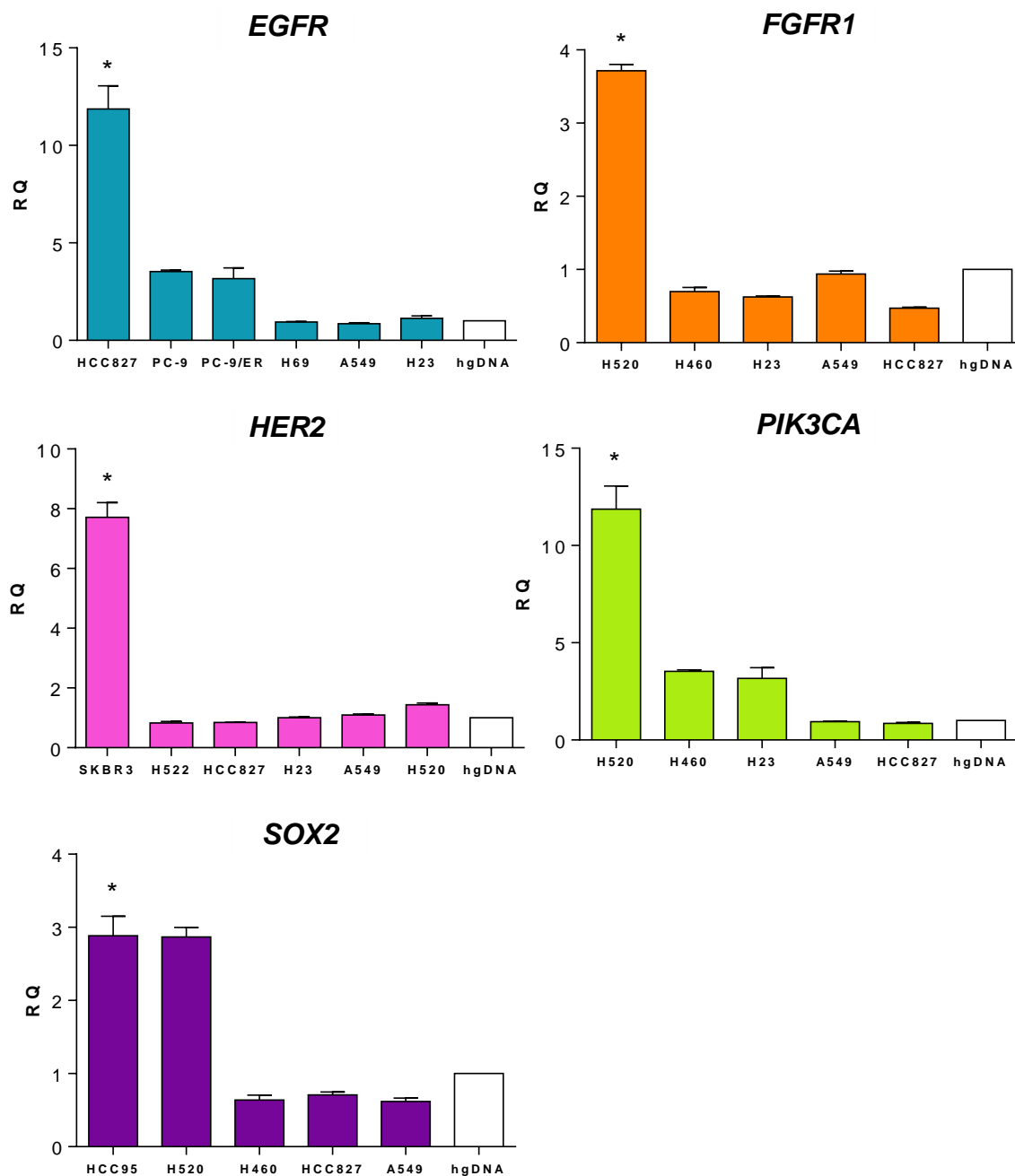
**Figure 5.2 – Cell line validation of qPCR assays for SCNA analysis**. Cancer cell lines of known positive amplification status (*) with respect to each target gene were used to validate each qPCR assay. Copy number was determined using the relative quantification (RQ or $\Delta\Delta C_T$) method using *RPPH1* as the endogenous reference gene and hgDNA as the calibrator (white). Columns represent the mean of three replicates and the positive error bars indicate SD.

There was an overall good agreement between copy number measurements obtained with the qPCR assays in Figure 5.2 and those reported in the literature for each cell line (Table 5.4). *EGFR* was amplified in cell line HCC827 at RQ 24, which is consistent with previous reports (RQ 20 - 36). Copy number of *PIK3CA* was higher than expected in cell line H520, but the previously reported amplification *PIK3CA* in this cell line used FISH rather than PCR. Conversely, *SOX2* copy number was low than expected in cell line HCC95, but the previously reported copy number was based on array CGH.

An RQ ≥ 2 (CN ≥ 4) has previously been considered to indicate amplification with respect to a specific locus (Kulka *et al.*, 2006; Suo *et al.*, 2004). In summary, all expected SCNAs were detected in the cell lines evaluated using RQ qPCR, suggesting that the assays could be used to evaluate copy number in DNA samples of unknown copy number status.

| Gene | Copy number | |
|------|------------|--------|
|  | Literature | Thesis |
| *EGFR* | 20-36 | 24 |
| *FGFR1* | 6-10 | 7.4 |
| *HER2* | 13-22 | 15.4 |
| *PIK3CA* | 5-10 | 26 |
| *SOX2* | 4.7 | 2.9 |

**Table 5.4 – Comparison of copy number measurements in cell lines previously reported to show amplification**.

## 5.3.2    Multiplex preamplification PCR

Multiplex preamplification is a preanalytical technique used to increase target abundance with minimal bias (Noutsias *et al.*, 2008; Korenkova *et al.*, 2015; Mengual *et al.*, 2008). Due to the low yield of cfDNA recovered from many plasma samples, it was necessary to use preamplification to enable multi-target qPCR analysis for SCNA. The approach was first modelled using cell line DNA (Fig 5.3).
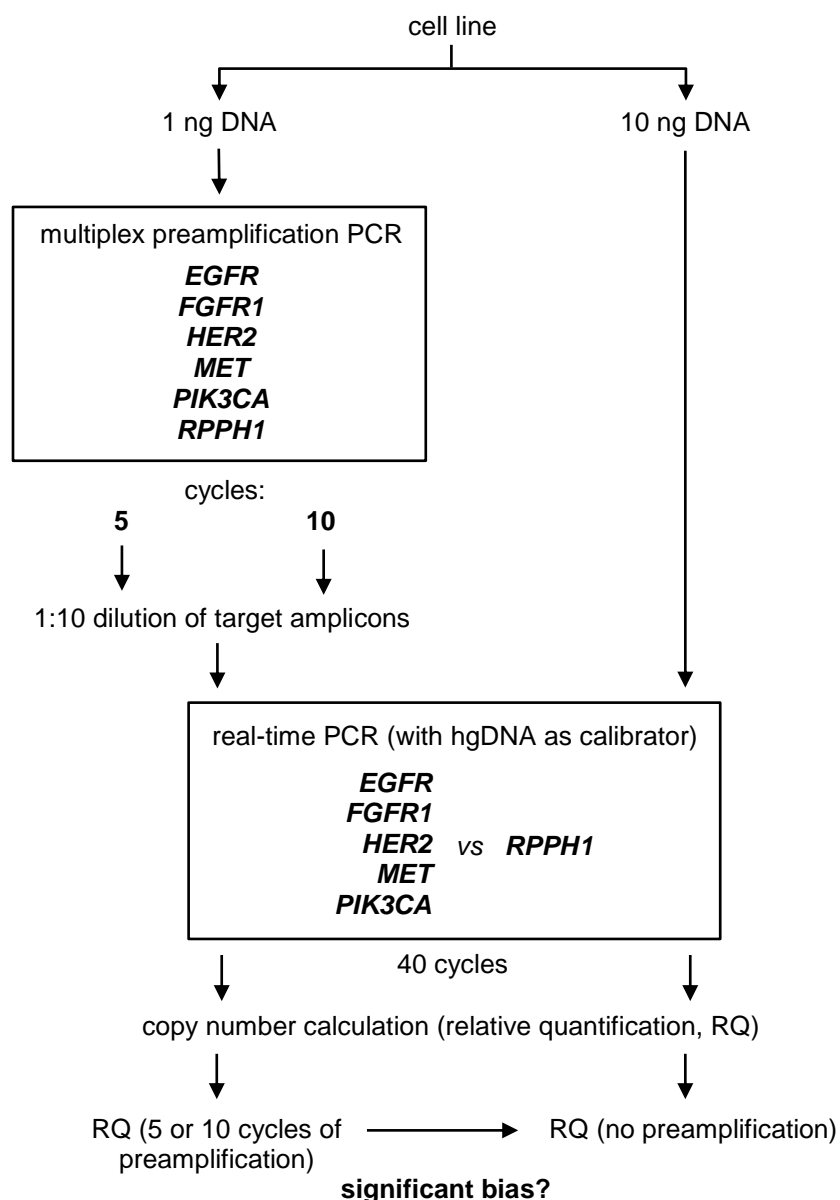


**Figure 5.3 – Schematic workflow of preamplification study.** SCNA was evaluated in cell lines using qPCR assays developed in Chapter 5.3.1, either using DNA directly after extraction or by first preamplifying the qPCR targets for 5 or 10 cycles. Copy number was calculated with the RQ method and compared for preamplified and non-preamplified DNA to assess for bias.

1 ng DNA (Table 5.3) was preamplified with a primer pool targeting *RPPH1*, *EGFR*, *FGFR1*, *HER2*, *MET* and *PIK3CA*. The assay targeting *SOX2* was designed after the validation process and hence was not included in the preamplification study. Preamplification was carried out comparing 5 and 10 cycles, and preamplified templates were diluted 1:10 to yield $C_T$ values in the range obtained with unamplified DNA. RQ values were calculated comparing unamplified and preamplified DNA after 5 and 10 cycles (Fig 5.3 and 5.4). Copy number measurements between preamplified (5X and 10X) and unamplified DNA were generally concordant in DNA that did not show amplification, whereas genes showing SCNA varied. 10 cycles of preamplification introduced significant bias to all assays ($P < 0.05$) compared to unamplified template, whereas all assays except *EGFR* were not significantly biased by 5 cycles of preamplification ($P > 0.05$). The low RQ of *PIK3CA* in H520 after 10 cycles of preamplification could have been caused by a technical error during preparation of the preamplification reaction. Based on these data, 5 cycles of preamplification was used for all subsequent experiments with cfDNA samples
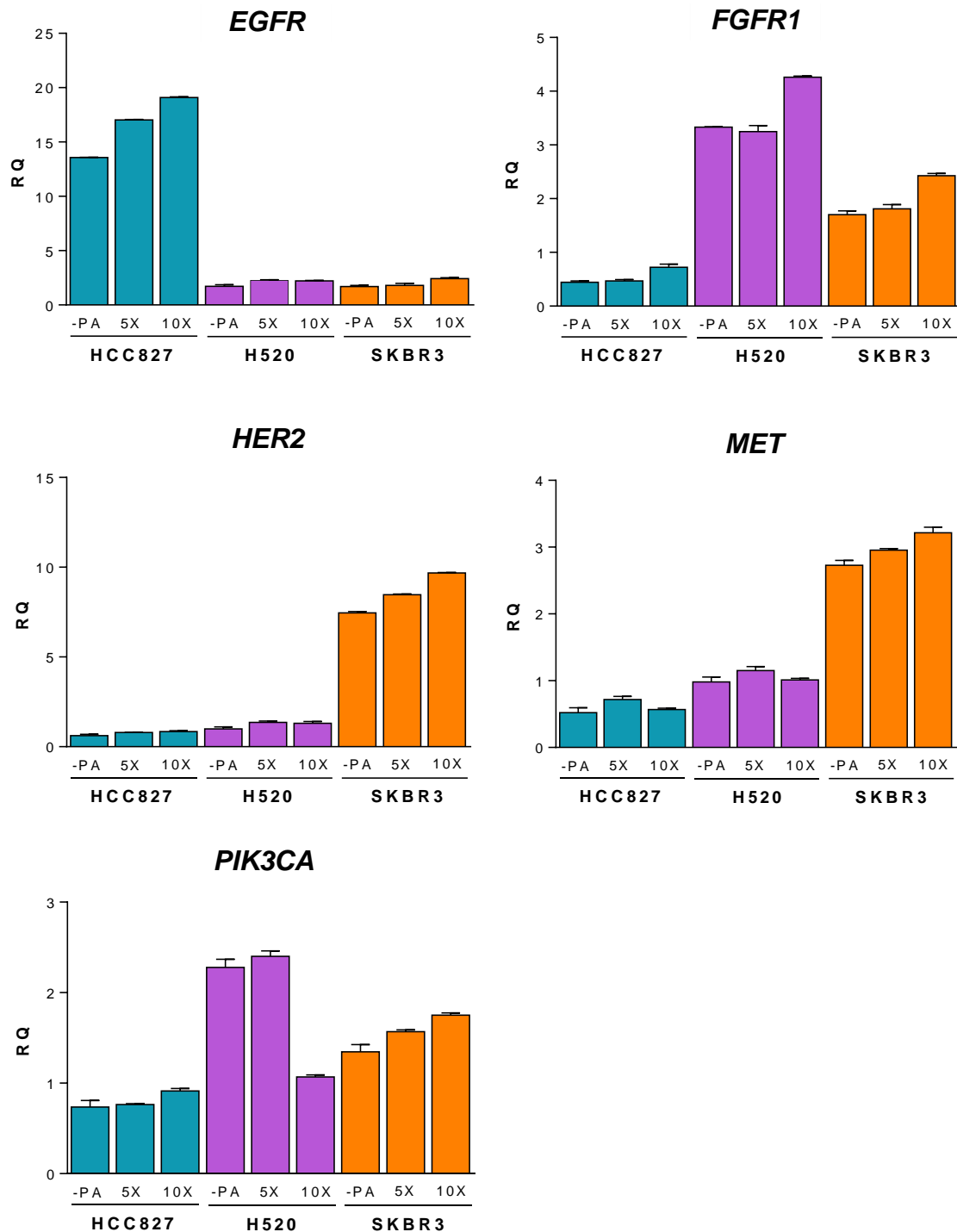
**Figure 5.4 – Comparison of gene amplification in amplified and unamplified cell line DNA by RQ qPCR.** Each qPCR assay was performed with 1 ng DNA extracted from three cell lines, before preamplification (-PA) and after 5 (5X) and 10 (10X) cycles of preamplification. Columns represent the mean of three replicates and the positive error bars indicate SD. RQ, relative quantification.

### 5.3.3    SCNA in locally advanced and advanced NSCLC

### 5.3.3.1    ReSoLuCENT cohort

SCNA was evaluated in cfDNA from a cohort of 40 patients with locally advanced or advanced NSCLC from the ReSoLuCENT study. These patients may have received any type of therapy for any number of cycles at the time of blood collection. Genomic DNA isolated from leukocytes was used as a germline control for each patient's cfDNA. In addition, cfDNA samples from 29 age and sex-matched healthy controls were evaluated to determine the range of SCNA values in healthy cfDNA. Clinicopathological characteristics of the patients in the ReSoLuCENT cohort are summarised in Table 5.5.

| Characteristic | *n* |
|---|---|
| *Gender* | |
| male | 22 |
| female | 18 |
| *Histology* | |
| AC | 12 |
| SqCC | 13 |
| NS | 15 |
| *Stage* | |
| IIIA | 10 |
| IIIB | 6 |
| IV | 24 |
| **Total** | **40** |

**Table 5.5 – Clinicopathological characteristics of patients enrolled on the ReSoLuCENT study.** Details of ethnicity and smoking history were not available. AC, adenocarcinoma; NS, not specified; SqCC, squamous cell carcinoma.

1 ng of cfDNA was preamplified for 5 cycles before RQ qPCR analysis of SCNA in 5 genes. Samples < 100 pg/μL were concentrated by lyophilisation prior to preamplification. 1 ng germline DNA was also preamplified and subject to qPCR in the same way. All 39 cfDNA samples produced signal at the qPCR stage,

indicating preamplification was successful. One germline DNA sample (P28) did not preamplify correctly, shown by the lack of qPCR signal across all assays.

Figure 5.5 shows the RQ of each gene in the 40 patient (●), 39 matched genomic DNA (○) and 29 control (∆) samples, calculated using the RQ method with human genomic DNA (98 genomes) used as the calibrator. RQ ≥ 2.5 and ≤ 0.5 were considered indicative of amplification and deletion, respectively, based on previous SCNA analyses using qPCR and accounting for a possible bias from the preamplification PCR (Page *et al.*, 2011; Lamy *et al.*, 2011; Bednarz *et al.*, 2009).

For *PIK3CA*, the RQ threshold was raised to 3.0, based on a wider range of RQ values obtained in both patients and healthy controls. There was no evidence of gene amplification in cfDNA in the healthy controls. Results in cfDNA showed a narrow range of RQ values (0.7 to 2.3) (Table 5.6), whereas RQ values ranged from 0.8 to 11.9 in cfDNA of patients with NSCLC (Table 5.7) .10/40 (25%) patient cfDNA samples showed at least one SCNA across the gene panel. 3 showed amplification at *EGFR*, 4 at *FGFR1*, 5 at *HER2*, 2 at *MET* and 8 at *PIK3CA* (Fig 5.11). No amplification was detected in germline DNA. Interestingly, half of all amplified samples showed co-amplification, with 2 samples harbouring amplification of all genes on the panel. Sample P08 exhibited high-level amplification of all genes, ranging from RQ 5.4 (*MET*) to 11.9 (*PIK3CA*). This patient had progressive stage IV disease and had a significantly higher cfDNA level than all other patients (298.5 ng/mL plasma vs. 7.4 ng/mL mean). All amplified samples were from patients with stage IV disease, except P30 (*PIK3CA*) and P38 (*HER2*), which were from stage IIIA. The mean RQ for each assay in patient cfDNA was significantly higher than in matched germline DNA ($P < 0.02$; FDR-adjusted significance ($q$) = 0.05). Plasma cfDNA levels were significantly elevated in samples with detectable SCNA relative to those without detectable SCNA (mean of plasma, 12.6 ng/mL vs 5.8 ng/mL, respectively) ($P = 0.007$; $q = 0.025$), and 8/10 patients with detectable SCNA had stage IV disease (vs. 16/30 without).

159
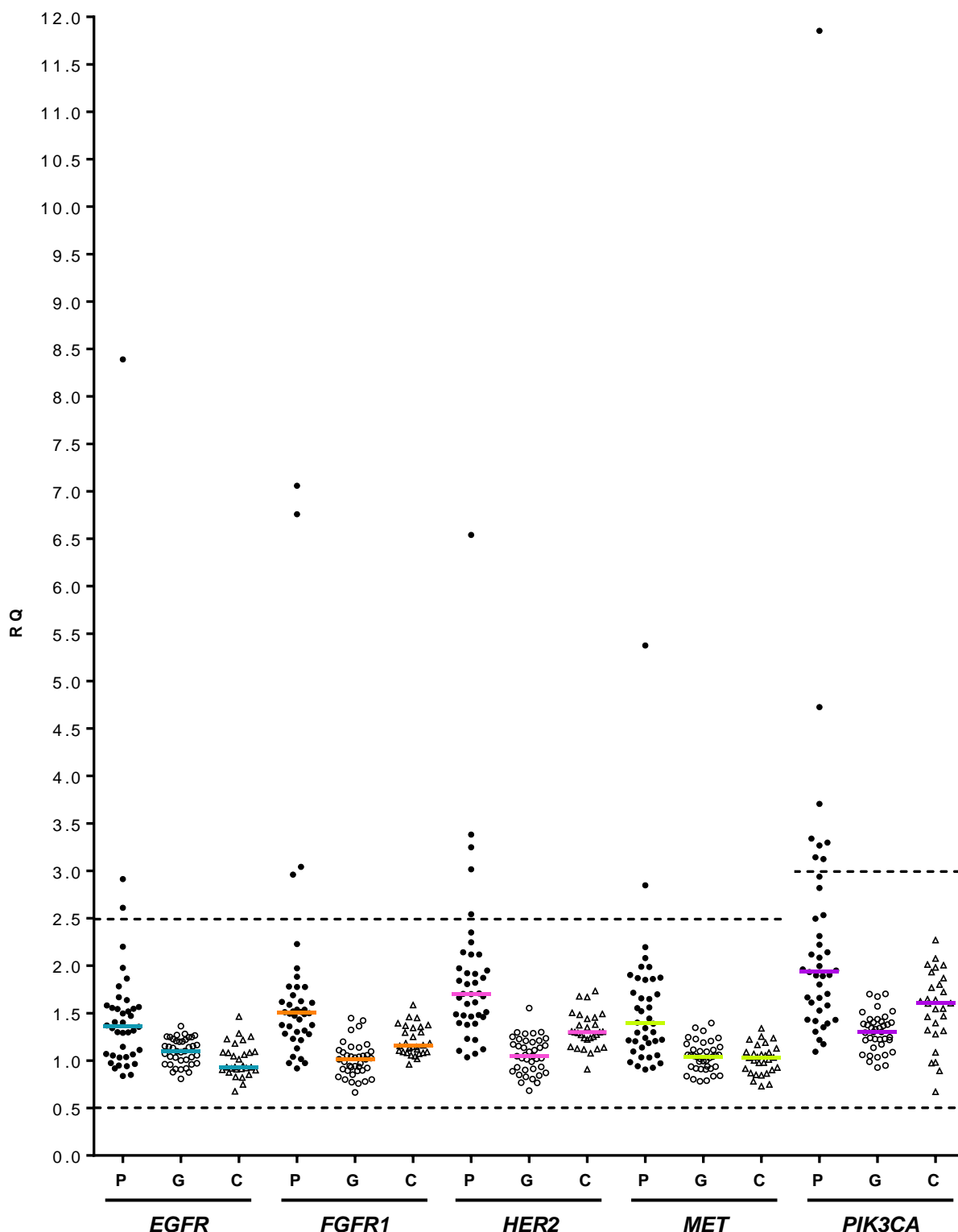
**Figure 5.5 – Analysis of SCNA in cfDNA and matched germline DNA of ReSoLuCENT patients by qPCR.** SCNA was evaluated in five genes from preamplified cfDNA (P; ●) and matched germline DNA (G; ○) of 40 patients with stage IIIA/IV NSCLC. cfDNA from 29 healthy controls (C; Δ) is also shown. Solid horizontal lines represent median values and perforated lines represent RQ thresholds. These data were subjected to multiple testing; therefore, *P* values were corrected for FDR according to the Benjamini-Hochberg procedure.

160

| ID | EGFR | FGFR1 | HER2 | MET | PIK3CA | ID | EGFR | FGFR1 | HER2 | MET | PIK3CA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| C01 | 1.23 | 1.25 | 1.67 | 1.34 | 1.09 | C16 | 0.93 | 1.34 | 1.38 | 0.98 | 1.40 |
| C02 | 0.94 | 1.05 | 1.44 | 0.75 | 1.80 | C17 | 0.91 | 1.09 | 1.24 | 0.98 | 2.08 |
| C03 | 1.06 | 1.37 | 1.37 | 1.25 | 1.47 | C18 | 0.91 | 1.15 | 0.91 | 0.87 | 0.67 |
| C04 | 0.92 | 1.08 | 1.49 | 0.85 | 1.98 | C19 | 0.92 | 1.18 | 1.30 | 1.09 | 1.77 |
| C05 | 1.08 | 1.46 | 1.30 | 1.17 | 1.54 | C20 | 1.25 | 1.39 | 1.27 | 1.24 | |
| C06 | 0.82 | 1.12 | 1.25 | 1.22 | 2.01 | C21 | 1.29 | 1.45 | 1.68 | 1.19 | 1.93 |
| C07 | 1.18 | 1.30 | 1.48 | 1.13 | 2.01 | C22 | 0.85 | 1.06 | 1.08 | 0.91 | 1.46 |
| C08 | 0.88 | 1.07 | 1.12 | 1.08 | 1.65 | C23 | 1.09 | 1.11 | 1.22 | 1.03 | 1.55 |
| C09 | 0.93 | 1.02 | 1.14 | 0.90 | 1.32 | C24 | 1.09 | 1.36 | 1.13 | 1.04 | 1.72 |
| C10 | 1.22 | 1.37 | 1.50 | 1.06 | 1.60 | C25 | 1.47 | 1.59 | 1.73 | 1.03 | 0.89 |
| C11 | 0.75 | 1.09 | 1.45 | 0.73 | 0.97 | C26 | 1.02 | 1.30 | 1.27 | 1.07 | 2.27 |
| C12 | 1.08 | 1.18 | 1.31 | 0.85 | 1.62 | C27 | 1.05 | 1.15 | 1.12 | 1.01 | 1.61 |
| C13 | 0.90 | 1.13 | 1.24 | 1.00 | 1.28 | C28 | 0.68 | 0.96 | 1.14 | 0.78 | 1.32 |
| C14 | 0.90 | 1.14 | 1.31 | 1.05 | 1.87 | C29 | 0.93 | 1.08 | 1.35 | 0.86 | 1.64 |
| C15 | 0.83 | 1.16 | 1.29 | 0.93 | 0.98 | | | | | | |

**Table 5.6 – RQ values per gene in healthy controls.**

| ID | EGFR | | FGFR1 | | HER2 | | MET | | PIK3CA | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Plasma | Leuk | Plasma | Leuk | Plasma | Leuk | Plasma | Leuk | Plasma | Leuk |
| P01 | 1.86 | 1.06 | 1.97 | 0.77 | 2.54 | 1.16 | 2.08 | 0.78 | 3.14 | 1.06 |
| P02 | 1.67 | 0.96 | 1.78 | 0.91 | 1.92 | 0.85 | 1.87 | 0.84 | 3.71 | 1.13 |
| P03 | 1.64 | 0.81 | 1.47 | 0.76 | 1.95 | 0.81 | 1.87 | 0.94 | 2.49 | 1.04 |
| P04 | 1.54 | 0.91 | 1.61 | 0.90 | 2.14 | 0.68 | 1.56 | 0.79 | 2.12 | 1.06 |
| P05 | 1.78 | 1.07 | 2.23 | 0.89 | 2.35 | 0.83 | 1.99 | 0.91 | 2.94 | 1.39 |
| P06 | 0.92 | 1.20 | 0.92 | 0.67 | 1.10 | 0.87 | 1.03 | 1.00 | 1.53 | 1.23 |
| P07 | 2.20 | 1.21 | 1.88 | 0.83 | 2.12 | 1.05 | 2.20 | 0.94 | 2.53 | 1.40 |
| P08 | 8.39 | 0.96 | 7.06 | 0.80 | 6.54 | 0.77 | 5.38 | 0.84 | 11.9 | 1.30 |
| P09 | 1.11 | 1.01 | 1.36 | 0.93 | 1.68 | 0.91 | 1.10 | 1.06 | 1.90 | 0.93 |
| P10 | 1.41 | 1.10 | 1.52 | 0.78 | 1.82 | 0.90 | 1.52 | 0.94 | 1.39 | 0.95 |
| P11 | 1.55 | 1.20 | 1.78 | 1.06 | 1.97 | 1.11 | 1.85 | 1.31 | 2.82 | 1.38 |
| P12 | 1.34 | 1.25 | 1.49 | 1.04 | 1.66 | 1.24 | 1.44 | 1.25 | 1.58 | 1.22 |
| P13 | 1.29 | 0.96 | 1.28 | 1.14 | 1.48 | 1.02 | 1.20 | 1.04 | 1.43 | 1.40 |
| P14 | 1.58 | 1.16 | 1.62 | 1.17 | 2.12 | 1.21 | 1.34 | 1.19 | 2.14 | 1.67 |
| P15 | 1.50 | 0.91 | 1.50 | 1.05 | 1.70 | 0.89 | 1.21 | 0.92 | 1.80 | 1.34 |
| P16 | 1.05 | 1.04 | 1.30 | 1.04 | 1.22 | 1.05 | 1.24 | 1.05 | 1.95 | 1.44 |
| P17 | 1.56 | 0.88 | 1.69 | 0.80 | 1.87 | 0.93 | 1.65 | 0.91 | 1.93 | 1.46 |
| P18 | 0.84 | 1.14 | 1.02 | 0.90 | 1.23 | 0.77 | 0.93 | 1.09 | 1.30 | 1.41 |
| P19 | 0.97 | 1.03 | 1.37 | 1.11 | 1.40 | 1.14 | 1.14 | 1.02 | 1.42 | 1.52 |
| P20 | 0.95 | 1.25 | 0.97 | 1.45 | 1.07 | 0.99 | 0.91 | 1.11 | 1.35 | 1.51 |
| P21 | 0.96 | 1.28 | 1.21 | 1.01 | 1.38 | 0.84 | 1.03 | 1.35 | 1.22 | 1.70 |
| P22 | 0.85 | 1.27 | 0.98 | 1.05 | 1.03 | 0.93 | 1.06 | 1.06 | 2.00 | 1.29 |
| P23 | 0.94 | 1.25 | 1.04 | 0.93 | 1.12 | 1.06 | 0.98 | 1.14 | 1.43 | 1.70 |
| P24 | 1.03 | 1.15 | 1.28 | 0.85 | 1.46 | 1.03 | 1.18 | 1.06 | 1.09 | 1.33 |
| P25 | 1.07 | 1.09 | 1.54 | 1.20 | 1.71 | 1.21 | 1.55 | 1.00 | 1.89 | 1.44 |
| P26 | 1.30 | 0.92 | 1.51 | 0.94 | 1.84 | 1.15 | 1.40 | 0.84 | 1.61 | 1.24 |
| P27 | 1.04 | 1.00 | 1.43 | 1.09 | 1.61 | 1.09 | 1.22 | 0.91 | 2.09 | 1.25 |
| P28 | 1.36 | | 1.37 | | 1.51 | | 1.90 | | 1.66 | |
| P29 | 1.47 | 1.14 | 1.54 | 1.03 | 1.70 | 1.14 | 1.72 | 1.40 | 3.27 | 1.23 |
| P30 | 1.52 | 1.25 | 1.63 | 1.10 | 1.39 | 1.19 | 1.70 | 1.06 | 3.34 | 1.17 |
| P31 | 2.61 | 1.36 | 3.04 | 1.32 | 3.25 | 1.23 | 1.86 | 1.04 | 1.96 | 1.30 |
| P32 | 1.37 | 1.27 | 6.76 | 1.42 | 1.46 | 1.29 | 1.30 | 1.10 | 3.12 | 1.22 |
| P33 | 1.40 | 0.97 | 1.59 | 1.03 | 1.92 | 1.30 | 1.39 | 1.24 | 1.90 | 1.37 |
| P34 | 2.91 | 1.15 | 2.96 | 1.15 | 3.02 | 1.17 | 2.85 | 1.12 | 4.72 | 1.09 |
| P35 | 1.98 | 1.22 | 1.78 | 0.92 | 2.25 | 1.30 | 1.99 | 1.18 | 2.31 | 1.34 |
| P36 | 1.32 | 1.13 | 1.50 | 1.01 | 1.81 | 1.28 | 1.66 | 1.21 | 2.22 | 1.38 |
| P37 | 1.15 | 1.04 | 1.23 | 0.94 | 1.60 | 1.02 | 0.97 | 0.99 | 1.17 | 1.05 |
| P38 | 1.07 | 0.88 | 1.32 | 0.97 | 3.38 | 1.01 | 0.94 | 0.80 | 1.67 | 0.99 |
| P39 | 1.30 | 1.10 | 1.13 | 0.99 | 1.49 | 1.25 | 1.21 | 1.03 | 3.30 | 1.22 |
| P40 | 1.56 | 1.23 | 1.51 | 1.36 | 1.48 | 1.55 | 1.30 | 1.23 | 1.70 | 1.57 |

**Table 5.7 – RQ values per gene in ReSoLuCENT patients.** Pink = samples subjected to NGS. Blue = amplification. Leuk, leukocyte.

20 cfDNA samples (pink ID in Table 5.7) from the ReSoLuCENT cohort were also sequenced with the Ion AmpliSeq™ Colon and Lung Cancer Research Panel v2 (Life Technologies) covering 22 genes (1850 hotspots), by Dr Barbara Ottolini (University of Leicester). The remaining 20 samples did not have sufficient (10 ng) cfDNA for sequencing. Variants with a coverage > 400x and a quality score > 25 were included in the analysis, and mispriming artefacts were filtered out using an algorithm developed by Dr Robert Hastings (Guttery *et al.*, 2015). Figure 5.6 shows all patients with detected SCNAs and/or mutations in cfDNA.
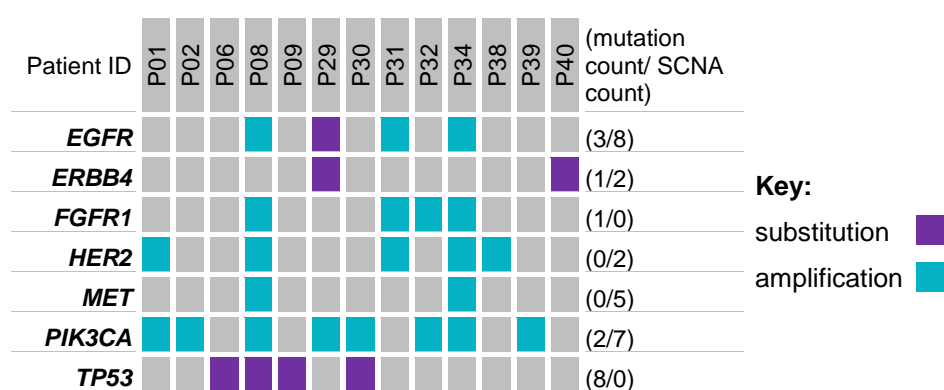


**Figure 5.6 – Oncoprint for cfDNA of ReSoLuCENT patients.** 40 cfDNA samples from patients with stage IIIA-IV NSCLC were sequenced with the Ion Torrent NGS platform, using the Ion AmpliSeq™ Colon and Lung Cancer Research Panel v2. Stringency parameters are detailed in the main text. SCNA was investigated using qPCR.

Variant allele frequency (VAF) ranged from 2.4 to 77.5% (mean, 13.6%) and was not correlated with original concentration of cfDNA ($R^2$ = 0.27). cfDNA levels were not significantly higher in patients with detectable mutations (mean of plasma, 11.5 ng/mL) than in patients without (mean of plasma, 11.8 ng/mL) ($P$ = 0.24)

Among the 20 samples, 6 showed evidence of circulating tumour DNA (ctDNA) by NGS after filtering (2 stage IIIA, 1 stage IIIB, 3 stage IV). Four samples had a single somatic mutation and 2 samples had 2 mutations. Interestingly, patient P08 with the highest cfDNA levels and amplification of all 5 genes also had a high frequency substitution mutation in *TP53* (c.A536T) at 77.5% VAF. Patient P29 showed amplification of *PIK3CA* (RQ, 3.3) and had a silent mutation in *EGFR* at

5.2% VAF. Patient P30 also had *PIK3CA* amplification and a dinucleotide change in *TP53* (c.CG743GA) at 3.5% VAF.

In summary, analysis of cfDNA in the ReSoLuCENT cohort (40 patients with stage IIIA-IV NSCLC) detected mutations in 6/20 patients by NGS (30% of cohort) and SCNA in 10/40 patients by qPCR (25%), with both mutations and SCNA detected in 3 patients. Combined analysis of mutations and SCNA identified ctDNA in a third (33%) of patients, despite no mutation analysis being performed in 20 patients. In addition, plasma cfDNA levels were significantly elevated in patients with SCNA ($P = 0.007$) but not mutated samples.

### 5.3.3.2    GALAXY-1 cohort

SCNA was evaluated in *EGFR*, *FGFR1*, *HER2*, *MET*, *PIK3CA* and *SOX2* in cfDNA of 143 patients from the GALAXY-1 clinical trial. Blood samples were obtained at baseline. All patients enrolled had received one prior systemic therapy, after which there was evidence of radiological progression (Ramalingam *et al.*, 2015). Clinicopathological characteristics of the GALAXY-1 cohort are summarised in Table 5.8 and detailed in Supplementary Table S1.1 (Appendix). Results were also compared with cfDNA from the 29 healthy controls (Table 5.6 in Chapter 5.3.3.1). Had matched germline (leukocyte) DNA samples in the patient cohort been available, these would also have been evaluated.

| Characteristic | *n* |
|---|---|
| *Ethnicity* | |
| Asian | 1 |
| Black | 1 |
| White | 141 |
| *Gender* | |
| male | 89 |
| female | 54 |
| *Histology* | |
| AC | 143 |
| *Smoking status* | |
| current | 37 |
| former | 73 |
| never | 33 |
| *Stage* | |
| IIIB | 7 |
| IV | 136 |
| **Total** | **143** |

**Table 5.8 – Clinicopathological characteristics of patients enrolled on the GALAXY-1 trial.** AC, adenocarcinoma.

1 ng cfDNA was preamplified for 5 cycles before RQ qPCR analysis in 143 patients from the GALAXY-1 trial.  Low concentration cfDNA samples (< 100 pg/µL) were concentrated by lyophilisation prior to preamplification. 16 patients were not included in the analysis due to failure (no signal) at the qPCR stage. qPCR was repeated on failed samples, which ruled out qPCR-specific error, but an additional 1 ng of cfDNA was not available to repeat the preamplification step in these failed samples. Failure of qPCR was not related to cfDNA concentration before preamplification (mean of 16 failed samples, 0.22 ng/µL; mean of 143 amplifiable samples, 0.25 ng/µL) There was no correlation between sample failure and disease stage or survival.

Figure 5.7 shows the RQ of each gene in the 143 patient (●) and 29 control (○) samples, calculated using the RQ ($\Delta\Delta C_T$) method with human genomic DNA as the calibrator. RQ values in cfDNA from 143 patients ranged from 0.2 to 4.7 (25.5 may be an outlier) (Table 5.9). There was no significant difference in total cfDNA levels between samples with and without detectable SCNA (mean 0.21 vs 0.25 ng/µL, respectively) ($P$ = 0.40). qPCR failed for a single assay in patient P136 (shown as diagonal lines in Tables). 18/143 (13%) patient cfDNA samples showed at least one SCNA across the gene panel. 7 showed amplification of *EGFR*, 2 of *FGFR1*, 3 of *HER2*, 4 of *MET* and 7 of *PIK3CA* (Fig 5.7). Interestingly, 4 of the 7 patients with *EGFR* amplification showed coamplification of *MET* and 1 patient sample showed coamplification of *FGFR1* and *HER2*.
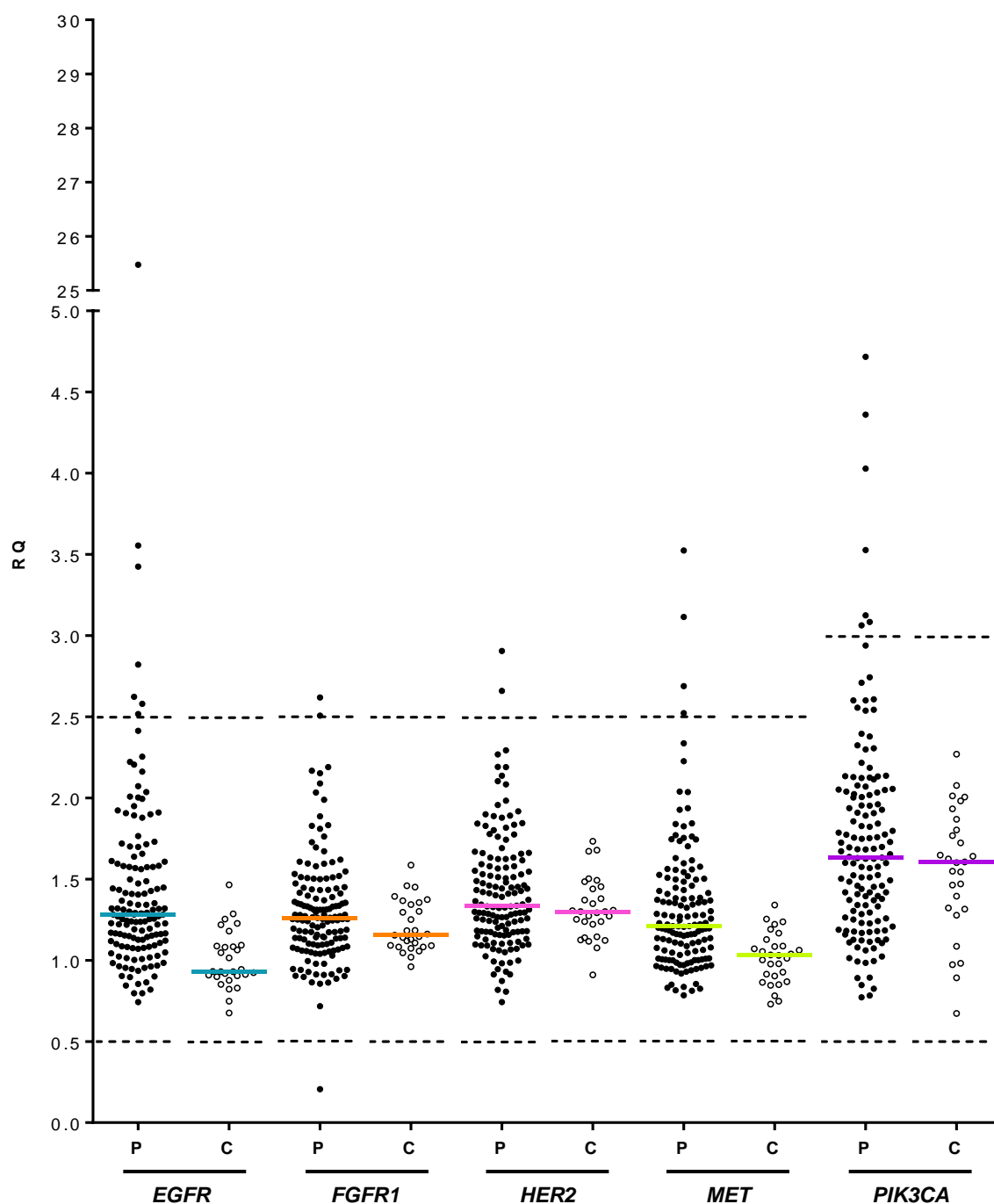
**Figure 5.7 – Analysis of SCNA in cfDNA of GALAXY-1 patients by qPCR.** SCNA was evaluated in five genes from preamplified cfDNA of 143 patients with stage IIIB/IV lung adenocarcinoma (P; ●) and 29 healthy controls (C; ○) using qPCR assays described in Chapter 5.3.1. Solid horizontal lines represent median values and perforated lines represent RQ thresholds.

| ID | EGFR | FGFR1 | HER2 | MET | PIK3CA | ID | EGFR | FGFR1 | HER2 | MET | PIK3CA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P001 | 1.58 | 1.35 | 1.57 | 1.58 | 2.07 | P073 | 1.22 | 1.50 | 1.58 | 1.19 | 1.86 |
| P002 | 2.16 | 0.94 | 1.11 | 1.70 | 1.26 | P074 | 1.12 | 1.83 | 1.92 | 1.38 | 1.78 |
| P003 | 1.28 | 0.89 | 0.82 | 1.83 | 1.68 | P075 | 0.97 | 1.34 | 1.29 | 1.23 | 1.00 |
| P004 | 1.22 | 1.31 | 1.50 | 1.41 | 1.65 | P076 | 1.28 | 1.27 | 1.33 | 1.38 | 2.54 |
| P005 | 3.55 | 1.08 | 1.45 | 3.11 | 2.30 | P077 | 1.50 | 1.51 | 1.80 | 0.93 | 2.02 |
| P006 | 2.52 | 0.98 | 1.11 | 2.04 | 2.32 | P078 | 0.86 | 0.94 | 1.15 | 1.06 | 1.12 |
| P007 | 1.61 | 1.51 | 1.84 | 1.60 | 1.50 | P079 | 1.10 | 1.32 | 1.51 | 0.93 | 1.51 |
| P008 | 2.82 | 1.09 | 1.10 | 1.41 | 1.76 | P080 | 1.22 | 1.18 | 1.24 | 1.00 | 1.82 |
| P009 | 1.66 | 1.26 | 1.28 | 1.54 | 1.40 | P081 | 1.29 | 0.86 | 1.83 | 1.13 | 1.18 |
| P010 | 1.56 | 1.03 | 1.07 | 1.19 | 1.67 | P082 | 0.80 | 1.20 | 1.07 | 1.17 | 1.01 |
| P011 | 2.25 | 1.37 | 1.67 | 2.04 | 2.60 | P083 | 0.95 | 1.28 | 1.45 | 1.30 | 2.03 |
| P012 | 1.04 | 1.22 | 0.98 | 0.99 | 1.37 | P084 | 1.02 | 1.43 | 1.46 | 1.00 | 1.69 |
| P013 | 1.28 | 1.24 | 1.09 | 1.16 | 1.51 | P085 | 25.48 | 0.91 | 1.26 | 3.52 | 1.58 |
| P014 | 1.70 | 1.38 | 1.17 | 1.43 | 1.95 | P086 | 1.27 | 1.14 | 1.19 | 1.84 | 1.84 |
| P015 | 1.90 | 1.31 | 1.18 | 1.38 | 1.18 | P087 | 3.42 | 1.31 | 1.53 | 2.52 | 2.61 |
| P016 | 1.41 | 1.55 | 1.18 | 1.16 | 1.45 | P088 | 1.25 | 1.67 | 1.74 | 1.23 | 2.13 |
| P017 | 1.31 | 1.28 | 1.33 | 1.01 | 1.12 | P089 | 0.95 | 1.16 | 1.13 | 1.00 | 1.91 |
| P018 | 1.57 | 1.25 | 1.10 | 1.45 | 1.62 | P090 | 0.95 | 1.26 | 1.63 | 1.05 | 2.02 |
| P019 | 1.13 | 1.22 | 1.16 | 1.21 | 2.31 | P091 | 1.09 | 1.50 | 1.53 | 1.22 | 2.56 |
| P020 | 1.73 | 1.99 | 2.66 | 1.76 | 1.17 | P092 | 0.98 | 1.07 | 1.25 | 0.96 | 1.06 |
| P021 | 1.99 | 1.59 | 1.88 | 1.50 | 1.18 | P093 | 1.15 | 1.24 | 1.45 | 0.92 | 1.46 |
| P022 | 1.95 | 1.61 | 2.29 | 1.74 | 1.23 | P094 | 0.90 | 1.08 | 1.34 | 0.83 | 0.89 |
| P023 | 1.77 | 0.95 | 1.20 | 1.52 | 0.85 | P095 | 1.44 | 1.06 | 1.62 | 1.14 | 1.42 |
| P024 | 1.41 | 1.32 | 1.10 | 0.95 | 0.99 | P096 | 0.80 | 1.00 | 1.16 | 0.84 | 1.21 |
| P025 | 2.21 | 1.31 | 1.44 | 1.16 | 1.79 | P097 | 2.04 | 1.43 | 2.19 | 1.36 | 1.93 |
| P026 | 1.24 | 1.04 | 0.91 | 0.95 | 1.25 | P098 | 1.32 | 2.51 | 1.90 | 1.37 | 1.89 |
| P027 | 1.08 | 0.86 | 0.91 | 0.96 | 1.34 | P099 | 0.96 | 1.25 | 1.26 | 1.1 | 1.44 |
| P028 | 1.11 | 0.86 | 0.81 | 1.12 | 1.28 | P100 | 1.02 | 1.31 | 1.16 | 1.21 | 1.42 |
| P029 | 1.61 | 1.60 | 1.63 | 1.63 | 2.19 | P101 | 1.29 | 1.45 | 1.78 | 1.08 | 1.94 |
| P030 | 1.71 | 1.60 | 1.78 | 2.23 | 1.29 | P102 | 1.13 | 1.27 | 1.31 | 0.85 | 1.11 |
| P031 | 1.17 | 1.33 | 1.41 | 1.21 | 1.73 | P103 | 0.90 | 1.10 | 1.17 | 1.00 | 1.33 |
| P032 | 1.58 | 2.15 | 1.98 | 1.56 | 2.54 | P104 | 0.82 | 1.11 | 1.30 | 0.81 | 1.70 |
| P033 | 1.15 | 1.14 | 1.18 | 1.01 | 1.57 | P105 | 1.05 | 1.27 | 1.44 | 0.85 | 2.05 |
| P034 | 1.08 | 0.90 | 1.40 | 0.94 | 1.29 | P106 | 1.35 | 1.17 | 1.29 | 1.22 | 1.42 |
| P035 | 1.19 | 1.39 | 1.43 | 1.28 | 1.95 | P107 | 1.23 | 1.26 | 1.69 | 1.28 | 1.73 |
| P036 | 0.90 | 1.28 | 1.29 | 0.97 | 1.02 | P108 | 1.29 | 0.21 | 1.09 | 1.18 | 1.49 |
| P037 | 1.01 | 1.10 | 1.29 | 0.98 | 2.05 | P109 | 1.16 | 1.34 | 1.89 | 1.09 | 4.72 |
| P038 | 0.85 | 1.25 | 1.27 | 1.25 | 1.63 | P110 | 1.43 | 1.13 | 0.93 | 1.37 | 1.23 |
| P039 | 1.31 | 1.81 | 1.82 | 1.94 | 1.53 | P111 | 1.57 | 1.09 | 1.32 | 1.36 | 1.18 |
| P040 | 1.88 | 2.19 | 2.14 | 1.84 | 3.06 | P112 | 1.34 | 1.35 | 1.47 | 1.26 | 1.41 |
| P041 | 1.26 | 1.25 | 1.55 | 1.15 | 1.60 | P113 | 1.41 | 1.15 | 1.26 | 1.36 | 1.52 |
| P042 | 1.09 | 1.17 | 1.37 | 1.01 | 1.63 | P114 | 1.32 | 0.91 | 1.13 | 1.06 | 1.40 |
| P043 | 0.96 | 1.20 | 1.27 | 1.05 | 2.38 | P115 | 1.24 | 1.83 | 2.08 | 1.75 | 3.13 |
| P044 | 0.98 | 1.28 | 1.33 | 0.95 | 1.70 | P116 | 1.14 | 1.06 | 1.18 | 1.11 | 2.04 |
| P045 | 1.30 | 1.58 | 1.33 | 1.27 | 1.75 | P117 | 2.62 | 1.17 | 1.55 | 1.62 | 2.71 |
| P046 | 1.16 | 1.42 | 1.41 | 1.00 | 2.74 | P118 | 0.84 | 1.10 | 1.34 | 0.99 | 2.13 |
| P047 | 1.03 | 1.35 | 1.41 | 0.95 | 1.46 | P119 | 1.01 | 1.26 | 1.36 | 0.97 | 1.11 |
| P048 | 1.49 | 1.34 | 1.66 | 1.03 | 1.64 | P120 | 0.94 | 1.05 | 1.28 | 1.00 | 2.06 |
| P049 | 1.15 | 1.25 | 1.58 | 1.06 | 1.77 | P121 | 0.74 | 0.91 | 1.09 | 0.79 | 1.08 |
| P050 | 1.24 | 1.44 | 1.46 | 1.18 | 2.60 | P122 | 1.00 | 0.94 | 1.24 | 0.82 | 2.13 |
| P051 | 2.58 | 1.41 | 1.57 | 2.69 | 1.60 | P123 | 1.24 | 1.50 | 1.24 | 1.20 | 1.30 |
| P052 | 1.57 | 1.14 | 1.43 | 1.21 | 1.38 | P124 | 1.45 | 1.44 | 1.52 | 1.32 | 2.11 |
| P053 | 1.91 | 1.53 | 1.83 | 1.75 | 1.24 | P125 | 2.22 | 1.89 | 2.10 | 2.34 | 4.36 |
| P054 | 1.43 | 1.33 | 0.95 | 1.27 | 1.17 | P126 | 1.41 | 1.07 | 1.25 | 1.36 | 2.01 |
| P055 | 1.37 | 1.14 | 1.11 | 1.20 | 0.98 | P127 | 1.20 | 1.47 | 1.66 | 1.13 | 1.57 |
| P056 | 1.91 | 0.91 | 1.66 | 1.46 | 0.89 | P128 | 1.08 | 1.36 | 1.34 | 0.97 | 1.19 |
| P057 | 1.70 | 1.03 | 1.29 | 1.27 | 1.48 | P129 | 1.27 | 1.69 | 1.62 | 1.14 | 2.40 |
| P058 | 1.14 | 1.25 | 1.21 | 0.95 | 1.80 | P130 | 1.23 | 1.51 | 1.89 | 1.56 | 1.96 |
| P059 | 1.17 | 1.18 | 1.23 | 0.82 | 1.83 | P131 | 1.64 | 1.62 | 1.96 | 1.50 | 2.00 |
| P060 | 1.59 | 1.05 | 1.29 | 1.11 | 2.12 | P132 | 1.35 | 2.09 | 1.85 | 1.50 | 1.16 |
| P061 | 1.71 | 1.45 | 1.62 | 1.29 | 2.22 | P133 | 2.00 | 2.17 | 2.19 | 1.43 | 1.37 |
| P062 | 1.07 | 0.90 | 0.87 | 1.20 | 1.91 | P134 | 1.92 | 2.03 | 1.51 | 1.77 | 1.11 |
| P063 | 1.72 | 1.04 | 1.49 | 1.16 | 3.08 | P135 | 1.34 | 1.40 | 1.48 | 1.16 | 0.78 |
| P064 | 1.08 | 1.13 | 1.42 | 1.35 | 1.13 | P136 | 1.25 | 1.10 | 0.74 |  | 4.03 |
| P065 | 1.44 | 1.18 | 0.99 | 1.53 | 1.07 | P137 | 1.17 | 1.73 | 1.18 | 1.39 | 2.13 |
| P066 | 1.58 | 0.72 | 1.00 | 1.56 | 0.77 | P138 | 1.32 | 1.33 | 1.39 | 1.51 | 1.75 |
| P067 | 2.07 | 0.93 | 1.07 | 1.28 | 1.13 | P139 | 1.32 | 1.25 | 1.17 | 1.34 | 2.08 |
| P068 | 1.47 | 1.05 | 1.06 | 1.08 | 1.26 | P140 | 1.18 | 1.14 | 1.05 | 1.22 | 1.68 |
| P069 | 1.89 | 1.52 | 1.76 | 1.34 | 3.53 | P141 | 2.41 | 1.19 | 1.59 | 1.73 | 1.60 |
| P070 | 1.29 | 1.06 | 1.15 | 1.19 | 1.34 | P142 | 1.24 | 1.14 | 0.99 | 1.11 | 2.14 |
| P071 | 1.12 | 2.62 | 2.90 | 1.01 | 1.63 | P143 | 1.16 | 0.98 | 1.03 | 1.48 | 1.75 |
| P072 | 2.01 | 1.76 | 2.27 | 1.93 | 2.94 |  |  |  |  |  |  |

**Table 5.9 – RQ values per gene in GALAXY-1 patients.** Blue = amplification. Yellow = deletion.

A disadvantage of qPCR for SCNA detection is that small differences in amplification efficiency of the target relative to reference can bias copy number measurements (Guescini *et al.*, 2008). Droplet digital PCR (ddPCR) may help to overcome this limitation, since the digital readout is not influenced by relative measurements. In addition, ddPCR can detect fine copy number differences and without the need for technical replicates. SCNA analysis using ddPCR often requires digestion of template DNA, since tandem repeats may be present in the same droplets, thus losing the ability to distinguish between one or multiple copies of a gene on a DNA strand. However, the fragmented nature of cfDNA obviates the need for this step. Samples showing amplification of *EGFR*, *MET* and *PIK3CA* by qPCR were evaluated by ddPCR, with each gene assayed in duplex with *RPPH1*. Borderline amplified samples were also run for each assay. Representative droplet clustering and fluorescence thresholds for *EGFR*, *MET*, *PIK3CA* and *RPPH1* are shown in Figure 5.8.
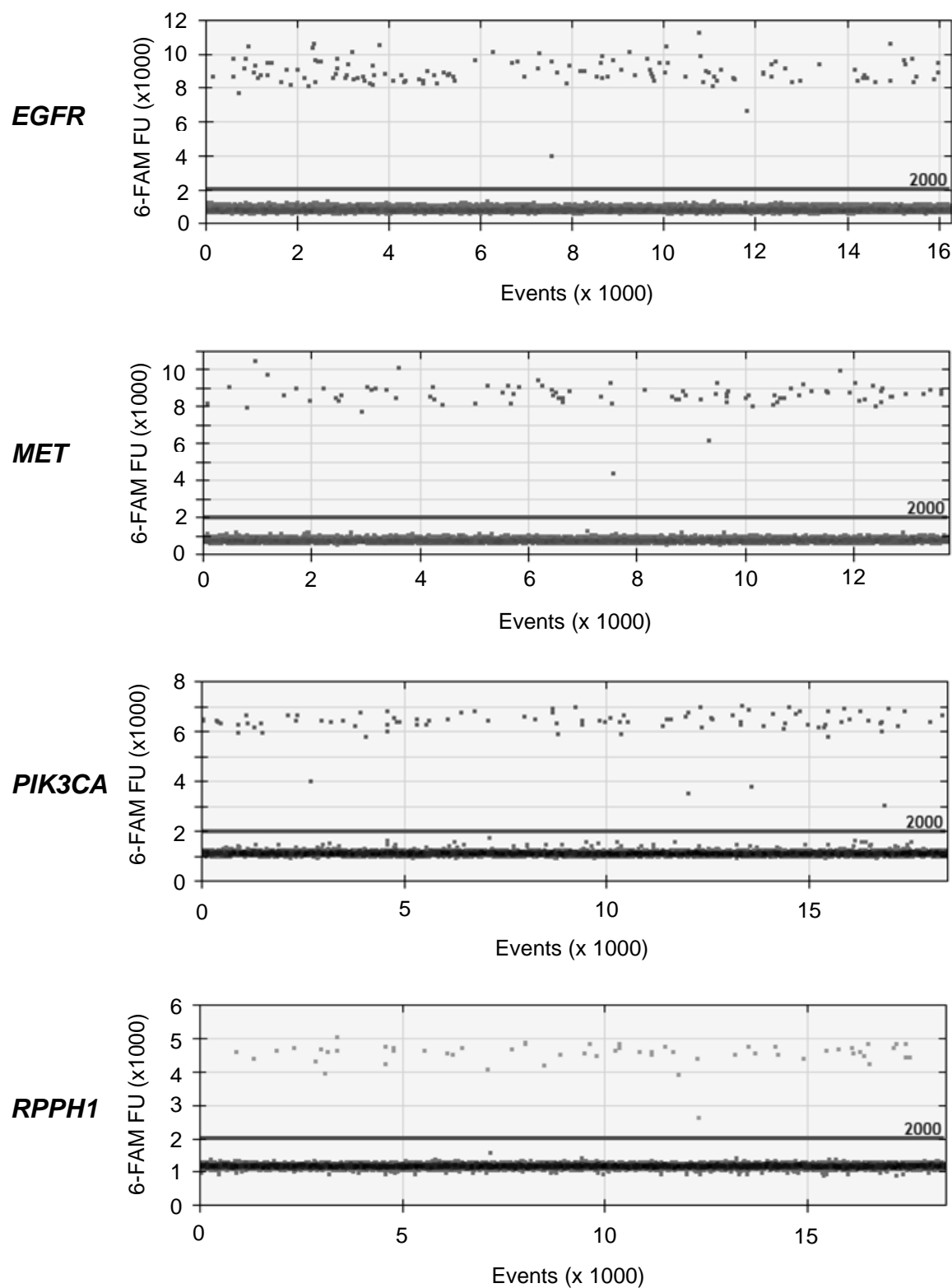
**Figure 5.8 – Droplet clustering in ddPCR assays for SCNA analysis and comparison with qPCR.** Each point represents a single droplet, which, in theory, contains a single copy of template. The height of each point is determined by the end-point fluorescence in the respective droplet, which is dependent on the number of cleaved 6-FAM™ fluorophores. FU, fluorescence units.

Paired qPCR and ddPCR copy number measurements for *EGFR*, *MET* and *PIK3CA* are shown in Table 5.10. The unit 'copy number ratio' (CNR) is used when comparing qPCR and ddPCR (instead of relative quantification), since ddPCR measurements are absolute values. For *EGFR*, 4/7 samples that were amplified by qPCR did not show amplification by ddPCR. Conversely, one sample was amplified by ddPCR only (qPCR CNR, 1.9, ddPCR CNR, 3.1). For *MET*, 4 samples were amplified by qPCR but not by ddPCR. However, 2 other samples showed amplification by ddPCR only (CNR, 2.7 and 2.6), the latter of which showed coamplification of *EGFR* by ddPCR. For *PIK3CA*, only two amplified samples by qPCR (CNR, 3.1 and 4.7) remained amplified by ddPCR (CNR, 9.0 and 4.0), the latter of which showed coamplification of *MET* by ddPCR. There was no significant difference in original cfDNA concentrations between amplified and unamplified samples by ddPCR (mean of eluates, 0.32 vs 0.25 ng/μL, respectively) (P = 0.43).

*SOX2* is located within the region of chromosome 3q that is frequently amplified in lung SqCC. Since this region also contains *PIK3CA*, it is possible that *SOX2* and *PIK3CA* may be coamplified in some cancers. To investigate this, cfDNA samples that were positive for *PIK3CA* amplification by *qPCR*, were assayed for *SOX2* amplification by ddPCR. No samples showed *SOX2* amplification (Table 5.11). This suggests a focal rather than chromosome-level amplification of *PIK3CA*.

Overall, a general decrease in CNR was observed with ddPCR relative to qPCR for all genes. The magnitude of the difference in copy number measurements was greater with qPCR than ddPCR, which could have been caused by standard curve-based artefacts in some assays; *PIK3CA* in particular (Table 5.2; 4% higher efficiency relative to *RPPH1*). Standard curves were constructed using high-integrity genomic DNA rather than cfDNA, and it is possible that some assays were more sensitive to DNA fragmentation (i.e. in cfDNA) than others, leading to a bias in copy number measurement.

| ID | EGFR | | ID | MET | | ID | PIK3CA | |
|---|---|---|---|---|---|---|---|---|
| | qPCR | ddPCR | | qPCR | ddPCR | | qPCR | ddPCR |
| P005 | 3.6 | 3.5 | P005 | 3.1 | 1.4 | P040 | 3.1 | 1.8 |
| P006 | 2.5 | 1.6 | P051 | 2.7 | 2.3 | P063 | 3.1 | 9.0 |
| P008 | 2.8 | 1.4 | P085 | 3.5 | 2.0 | P069 | 3.5 | 2.1 |
| P051 | 2.6 | 1.8 | P004 | 1.4 | 1.1 | P109 | 4.7 | 4.0 |
| P085 | 25.5 | 1.3 | P006 | 2.0 | 1.2 | P115 | 3.1 | 2.3 |
| P087 | 3.4 | 2.6 | P007 | 1.6 | 1.0 | P125 | 4.4 | 2.3 |
| P117 | 2.6 | 2.8 | P008 | 1.4 | 0.7 | P136 | 4.0 | 1.7 |
| P004 | 1.2 | 1.2 | P009 | 1.5 | 1.5 | P002 | 1.3 | 1.4 |
| P007 | 1.6 | 1.2 | P011 | 2.0 | 1.5 | P011 | 2.6 | 0.6 |
| P009 | 1.7 | 1.2 | P069 | 1.3 | 1.2 | P050 | 2.6 | 1.8 |
| P011 | 2.3 | 2.2 | P072 | 1.9 | 1.7 | P065 | 1.1 | 0.6 |
| P069 | 1.9 | 3.1 | P109 | 1.1 | 2.7 | P066 | 0.8 | 0.8 |
| P072 | 2.0 | 2.1 | P117 | 1.6 | 2.6 | P067 | 1.1 | 2.7 |
| | | | | | | P068 | 1.3 | 1.4 |
| | | | | | | P072 | 2.9 | 1.5 |
| | | | | | | P076 | 2.5 | 1.1 |
| | | | | | | P087 | 2.6 | 1.7 |
| | | | | | | P117 | 2.7 | 1.4 |
| | | | | | | P132 | 1.2 | 2.1 |
| | | | | | | P134 | 1.1 | 2.1 |
| | | | | | | P135 | 0.8 | 1.0 |

**Table 5.10 – Comparison of qPCR and ddPCR data for SCNA analysis in cfDNA of patients with advanced lung adenocarcinoma.** Genomic amplifications previously identified by qPCR in *EGFR*, *MET* and *PIK3CA* were assessed by ddPCR. Randomly selected cfDNA samples that did not show amplification of these genes by qPCR were also assayed. Blue = amplification. Values are copy number ratios (target vs *RPPH1*).

| ID | *PIK3CA* (ddPCR) | *SOX2* (ddPCR) |
|---|---|---|
| P040 | 1.8 | 1.1 |
| P063 | 9.0 | 1.5 |
| P069 | 2.1 | 1.7 |
| P109 | 4.0 | 1.3 |
| P115 | 2.3 | 1.3 |
| P125 | 2.3 | 1.4 |
| P136 | 1.7 | 1.3 |

**Table 5.11 – ddPCR analysis of *SOX2* copy number in *PIK3CA*-amplified cfDNA samples.** Copy number of *PIK3CA* and *SOX2* determined by ddPCR. Values are copy number ratios (target vs *RPPH1*). Blue = amplification.

### 5.3.3.2.1    Correlation of SCNA in cfDNA with survival

Survival characteristics were available for all GALAXY-1 patients. Cox regression analysis revealed no significant decrease in PFS (HR, 1.30; 95% CI, 0.75 – 2.24; $P$ = 0.27) but a significant decrease in OS (HR, 1.78; 95% CI, 1.05 – 3.02; $P$ = 0.03) (Fig 5.9) in patients with SCNA detected by qPCR and/or ddPCR (n = 18).



**Figure 5.9 – Kaplan-Meier survival curves based on SCNA status.** Comparison of (A) progression-free survival and (B) overall survival between patients with (n = 18) and without (n = 125) SCNA detected in cfDNA. PFS, progression-free survival.

### 5.3.3.2.2 Correlation of mutations in cfDNA with survival

cfDNA from the GALAXY-1 patient cohort was also deep-sequenced with the Ion Torrent NGS platform using the Ion AmpliSeq™ Cancer Hotspot Panel v2 (Life Technologies) covering 50 genes, by Mr C. Rakhit (MRC Toxicology Unit, Leicester). Variants with > 400x coverage and quality > 25 were included, and mispriming artefacts were filtered out using an algorithm designed by Dr Robert Hastings (Guttery *et al.*, 2015). Mutations and SCNA are shown in Figure 5.10.

Variant allele frequency (VAF) ranged from 3.6 to 39.3% (mean, 14.7%) and was not correlated with total cfDNA levels ($R^2$ = 0.0016) or the total number of mutations detected per sample ($R^2$ = 0.02). However, cfDNA levels were significantly higher in patients with detectable mutations (mean of eluate, 0.43 ng/μL) than in patients without (mean of eluate, 0.21 ng/μL) (*P* < 0.0001).

Among the 143 samples, 28 showed evidence of ctDNA by NGS after filtering. 18 samples had a single somatic mutation (17 with a substitution and 1 with a deletion) and 10 samples had multiple mutations. Patients with mutations had a significantly shorter PFS (HR, 1.61; 95% CI, 0.99 – 2.62; *P* = 0.02) and OS (HR, 1.65; 95% CI, 1.02 – 2.68; *P* = 0.04) than patients with no detectable mutations in cfDNA by Cox regression analysis (Fig 5.11). The 18 samples with detected SCNA by qPCR and/or ddPCR, 2 harboured mutations, and combined analysis of mutations and SCNA (by both PCR methods) detected alterations in 44 patients (31%). Patients with mutations and/or SCNA had a significantly shorter PFS (HR, 1.50; 95% CI, 1.01 – 2.21; *P* = 0.02) and OS (HR, 1.88; 95% CI, 1.24 – 2.84; *P* = 0.003) than patients with no alterations in cfDNA (Fig 5.12).

Patient P040 had a substitution in *RB1* (c.C1666T) at 8.9% VAF and amplification of *PIK3CA* (CNR, 3.1 by qPCR and 1.8 by ddPCR) and *MET* (CNR, 1.1 by qPCR and 2.7 by ddPCR). This patient had stage IV disease and a PFS of 21 days. Patient P109 had a substitution in *HRAS* (c.T81C) at the highest VAF in the cohort (39.3%) and amplification of *PIK3CA* by both qPCR and ddPCR (CNR, 4.7 and 4.0, respectively). This patient had stage IV disease and poor PFS (22 days on docetaxel vs. mean of 137 days for stage IV patients on docetaxel (*n* = 69)).
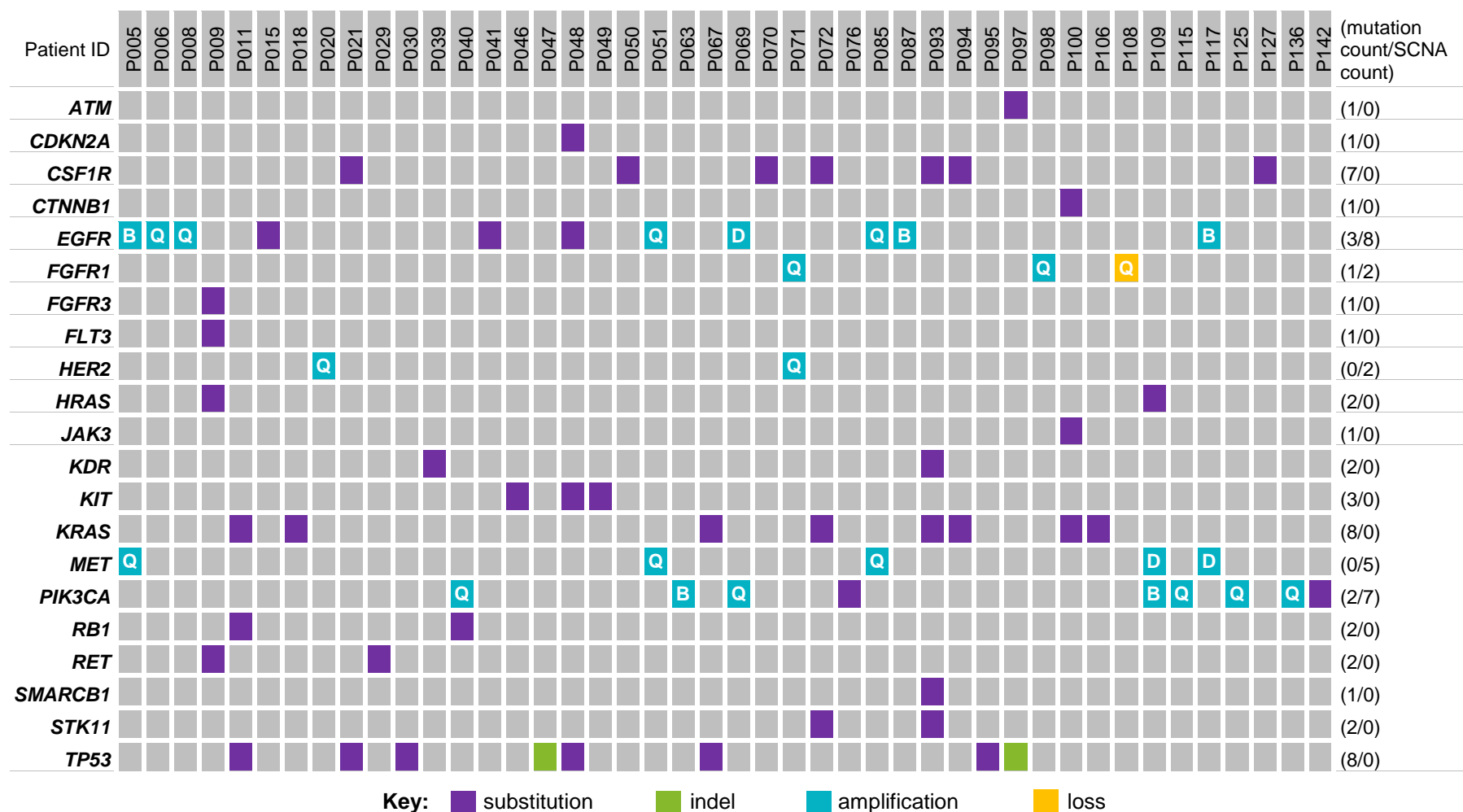
**Figure 5.10 – Oncoprint for cfDNA of GALAXY-1 patients.** 143 cfDNA samples from patients with advanced lung adenocarcinoma were sequenced with the Ion Torrent NGS platform, using the Ion AmpliSeq™ Cancer Hotspot Panel v2. Stringency parameters are detailed in the main text. SCNA was investigated using qPCR (Q) and ddPCR (D). B, amplified by both qPCR and ddPCR. Only patients with detected mutations and/or SCNA are shown.
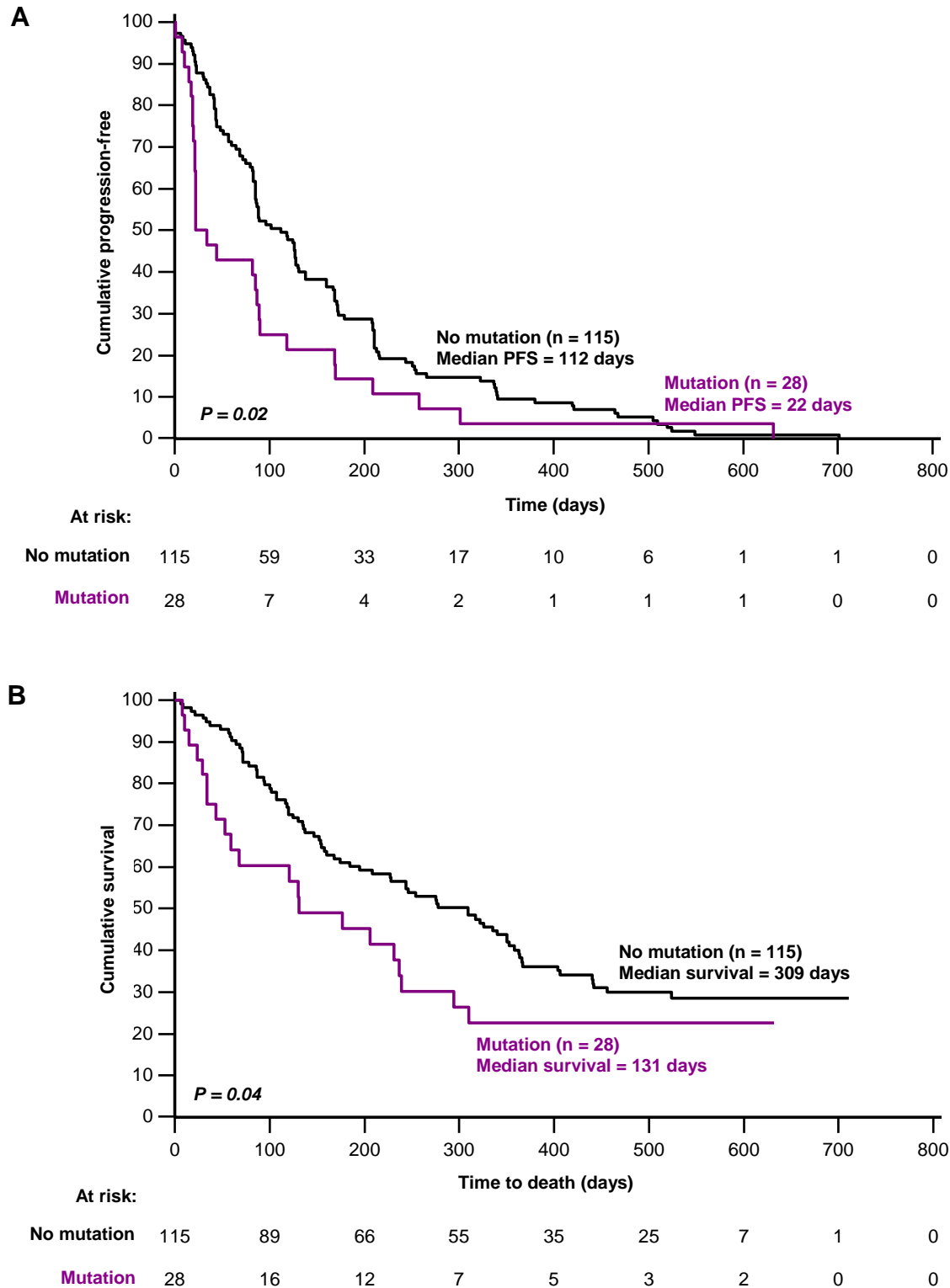
**Figure 5.11 – Kaplan-Meier survival curves based on mutation status.** Comparison of (A) progression-free survival and (B) overall survival between patients with (n = 28) and without (n = 115) one or more mutations detected in cfDNA. PFS, progression-free survival.
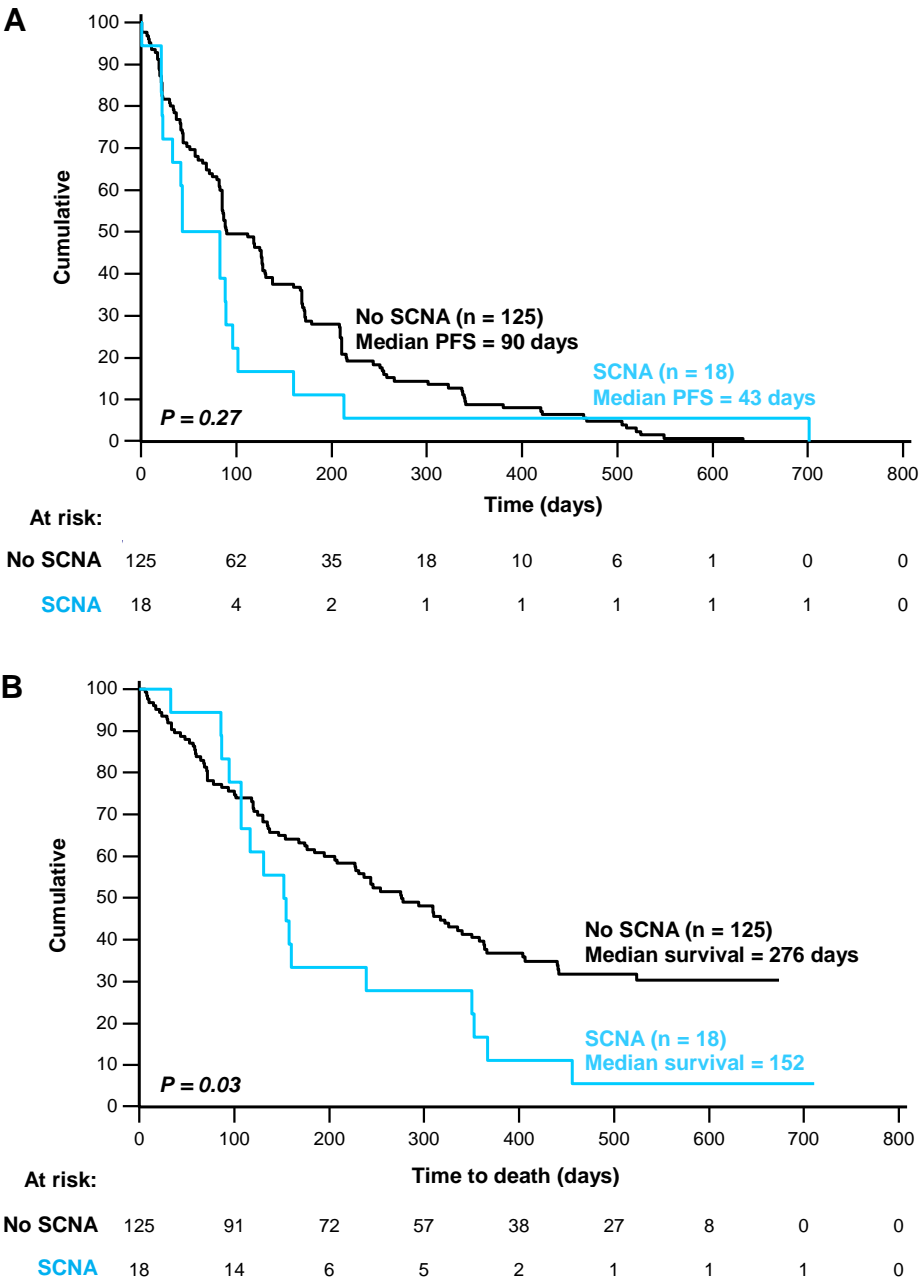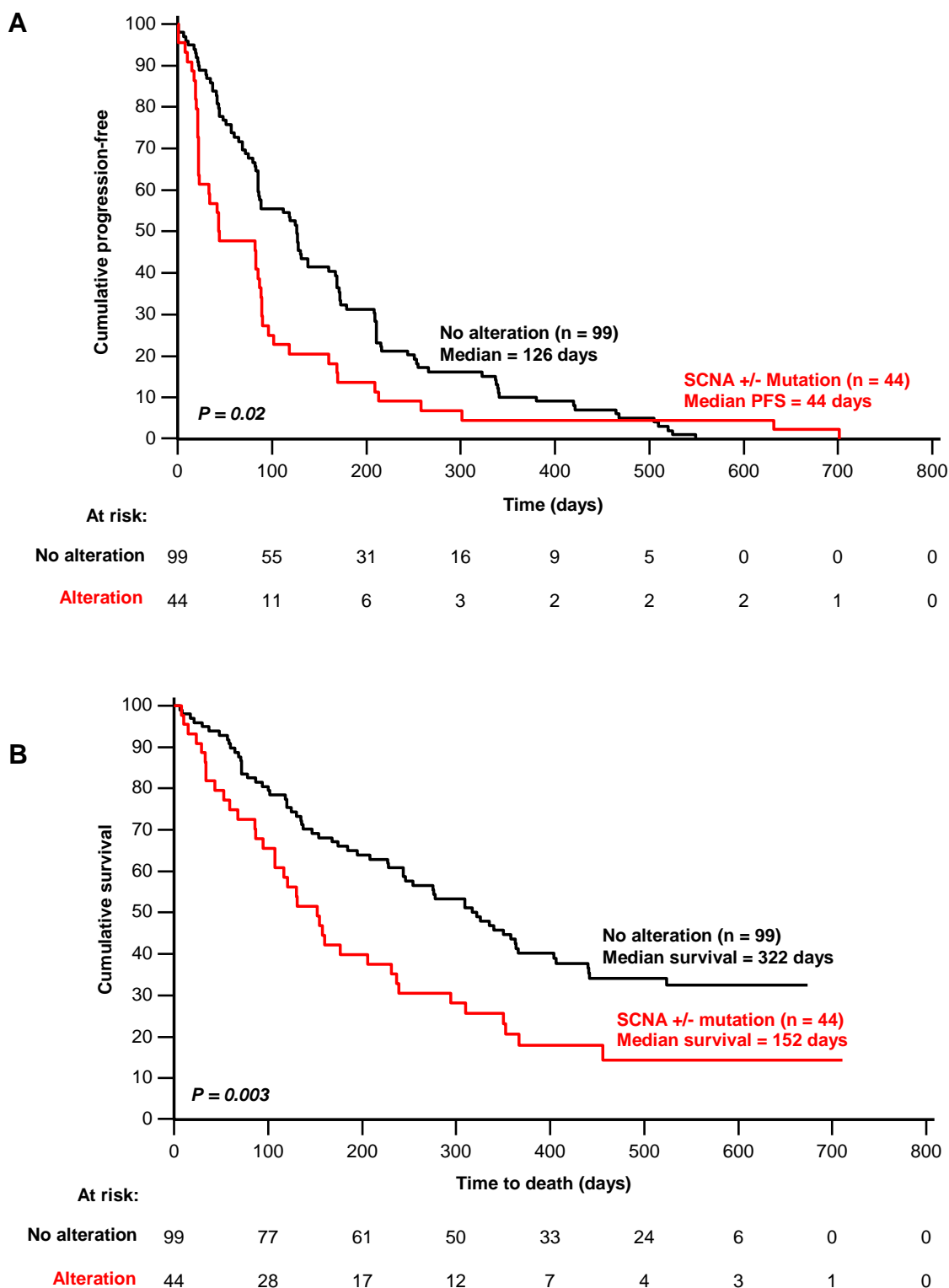
**Figure 5.12 – Kaplan-Meier survival curves based on SCNA and mutation status.** Comparison of (A) progression-free survival and (B) overall survival between patients with (n = 44) and without (n = 99) one or more SCNA and/or mutations detected in cfDNA. PFS, progression-free survival.

Since mutation and/or SCNA status were significantly correlated with survival outcomes in the GALAXY-1 cohort, patients were sub-grouped into the treatment arms (docetaxel alone and docetaxel plus ganetespib) to investigate whether treatment further influenced survival outcomes. Using Cox regression analysis, there was no significant influence of treatment on survival in patients grouped by SCNA or combined SCNA and mutation status. In patients with no detected mutations (n = 115), there was a significant decrease in PFS in the docetaxel arm relative to the docetaxel plus ganetespib arm (HR, 1.56; 95% CI, 0.99 - 2.43; $P$ = 0.05) and a favourable trend towards poor OS in these patients (HR, 1.4; 95% CI, 0.94 - 2.29; $P$ = 0.09) (Rakhit *et al.*, 2016 [submitted]).

In summary, analysis of cfDNA in the GALAXY-1 cohort (143 patients with advanced lung adenocarcinoma) by NGS detected mutations in 28 patients (19% of cohort), and these patients had significantly higher total plasma cfDNA levels than patients with no detectable mutations ($P$ < 0.0001). Addition of a five-gene panel for SCNA analysis by PCR provided evidence of SCNA in a further 16 patients. Thus, combined analysis of mutations and SCNAs identified ctDNA in a total of 44 patients (31% of cohort). The subset of patients with SCNA (n = 18) had significantly poorer OS than patients with no detectable SCNA ($P$ = 0.03), and this survival trend was seen albeit with lesser significance when stratifying patients based on mutation status (one or more mutations vs. no mutations), in addition to a significant decrease in PFS. Importantly, the difference in survival outcomes was most significant when SCNA and mutation statuses were combined ($P$ = 0.02 and $P$ = 0.003 for PFS and OS, respectively). Thus, analysis of SCNA in the GALAXY-1 cohort increased the ability to stratify patients based on survival outcomes compared with NGS alone.

### 5.3.3.3 Summary of ReSoLuCENT and GALAXY-1 cohort results

The ReSoLuCENT and GALAXY-1 cohorts comprised a total of 183 patients with locally advanced and advanced NSCLC. Patients either had evidence of progression after one systemic therapy (GALAXY-1) or were undergoing treatment (ReSoLuCENT). All patients' cfDNA samples were analysed for SCNA using and panel of five genes, and a total of 28 samples (15%) showed evidence of amplification or deletion. In the ReSoLuCENT cohort, patients with SCNA had significantly higher plasma cfDNA levels than patients with no evidence of ctDNA ($P = 0.007$). In the GALAXY-1 cohort, this held true for patients with mutations (but not SCNA) ($P < 0.0001$). Half of the ReSoLuCENT samples (20/40) and all of the GALAXY-1 samples were subjected to NGS: 21% (34/163) of patients were positive for at least one mutation. The mutation VAFs ranged from to 2.4 to 77.5%, and two samples with high VAF (*HRAS* at 39.3% in P109 and *TP53* at 77.5% in P08) were associated with gene amplification.

Of particular interest was the strong association between cfDNA alterations and survival outcomes (OS and PFS) in the GALAXY-1 cohort. Analysis of SCNA by qPCR and ddPCR identified 18 patients with evidence of ctDNA, and these patients had significantly poorer OS (HR, 1.78; $P = 0.03$) than patients with no detectable SCNA. NGS identified a further 26 patients with significantly worse PFS (HR, 1.78; $P = 0.02$) and OS (HR, 1.65; $P = 0.04$) than patients with no detectable mutations. Combined analysis of mutations and SCNA identified a subset of 44 patients with significantly poorer PFS (HR, 1.50; $P = 0.02$) and OS (HR, 1.88; $P = 0.003$) when compared to patients with no detectable alterations in cfDNA, and this was more significant for OS than with analysis of SCNA or mutations alone. Thus, SCNA is an important alteration in cfDNA that identified a subset of patients in this GALAXY-1 cohort with significantly poorer survival. PCR-based SCNA panels assessing commonly altered genes may have clinical utility for the evaluation of cfDNA, owing to their low cost (relative to NGS), simplicity of workflow and high throughput.

## 5.4        Discussion

The aim of this chapter was to determine whether SCNA in genes commonly amplified in NSCLC can be detected in the cfDNA of patients with locally advanced and advanced disease. A panel of genes (*EGFR*, *FGFR1*, *HER2*, *MET* and *PIK3CA*) was analysed by qPCR following multiplex preamplification and some SCNAs were validated with ddPCR. Where possible, copy number data was combined with Ion Torrent™ NGS mutation data from the same patients and correlated with survival outcomes.

### 5.4.1        Methodological limitations

The low concentration of cfDNA obtained from the blood samples in this study necessitated use of multiplex preamplification, enabling multi-target analysis by PCR. While several studies have reported the successful use of preamplification, others have reported its potential to introduce significant experimental bias in copy number measurements (Noutsias *et al.*, 2008; Korenkova *et al.*, 2015; Mengual *et al.*, 2008; Sanders *et al.*, 2011). The performance of all assays except *EGFR* was not significantly influenced by preamplification, and bias in this assay was only observed in the highly *EGFR*-amplified cell line, HCC827. Concordant with this, assays generally showed greater bias in cell lines harbouring high-level amplifications. In the ReSoLuCENT cohort, the mean copy numbers of all assays except *PIK3CA* in germline DNA samples was around 1.0, which confirms no significant bias from preamplification at the patient analysis stage.

*PIK3CA* copy numbers were more dispersed in both patient cohorts, and the mean RQ was greater than that of other genes in germline DNA of the ReSoLuCENT cohort, likely due to a 4% difference in PCR efficiency between *PIK3CA* and *RPPH1*. This observation highlights an inherent pitfall of qPCR for analysis of copy number, since qPCR assays are rarely 100% efficient and PCR inhibitors can have an assay-specific influence on performance (Kontanis and Reed, 2006). It has been reported that a 4% difference in efficiency of amplification between target and reference can bias the $\Delta C_T$ value up to 400%

(Guescini *et al.*, 2008). For this reason, the RQ threshold for amplification of *PIK3CA* was raised to 3.0. The paralogue ratio test (PRT) is reported to be more accurate than qPCR for copy number calculation, since a copy-neutral region of the genome is simultaneously amplified with the region of interest using the same primer pair  (Armour *et al.*, 2007; Aldhous *et al.*, 2010). However, the PRT approach could not be adopted in this project because copy-neutral regions of the genome that share sequence homology to the genes of interest could not be identified.

## 5.4.2     SCNA analysis in cfDNA of patients with NSCLC

cfDNA from a total of 183 patients with locally advanced or advanced NSCLC were evaluated for somatic copy number alteration (SCNA) in this Chapter. qPCR analysis detected SCNA in 15% of patients, which were all amplification events, except one single deletion event (*FGFR1*). Interestingly, the frequency of coamplification in amplified samples was higher than anticipated, with a third of amplified samples across both cohorts exhibiting coamplification. *MET* and *EGFR* (both on chromosome 7) are frequently coamplified in NSCLC (20 - 30% of *EGFR*-amplified tumours), and coamplification of *MET* was frequently observed in this study. (Cappuzzo *et al.*, 2009b; Beau-Faller *et al.*, 2008). However, other observed coamplifications, such as *HER2/FGFR1* and *HER2/PIK3CA*, have not been reported.

In the GALAXY-1 cohort, ddPCR was used to evaluate copy number in all samples with amplification by qPCR, since it is not reliant on equal efficiency of target and reference assays. In line with previous reports, there was a general downward shift in copy number measurements by qPCR relative to ddPCR, possibly due to an overestimation by qPCR (Whale *et al.*, 2012; Devonshire *et al.*, 2014; Bharuthram *et al.*, 2014). However, by ddPCR there was an overall tendency of samples that reached amplification by qPCR to retain a higher copy number than those not amplified by qPCR, showing the ability of qPCR to distinguish between higher and lower copy genes. Interestingly, GALAXY-1

patients with mutations had significantly higher cfDNA levels, whereas this was true for ReSoLuCENT patients with SCNA.

Half of the cfDNA samples in the ReSoLuCENT cohort and all samples in the GALAXY-1 patient cohort were sequenced by Ion Torrent™ NGS. Combined SCNA and NGS analysis identified evidence of ctDNA in a third of patients from both cohorts. Although only few samples exhibited both SCNA and mutations, the sequencing panels covered only 22 (ReSoLuCENT study) or 50 (GALAXY-1 study) genes, and the SCNA panel only 5 genes. Therefore, speculation that individual tumours may have been predominantly mutation- or SCNA-driven should be avoided.

In some patients, interesting correlations were made between mutation/CNA data and clinicopathological characteristics. For example, a patient with a high frequency mutation in *HRAS* and amplification of *PIK3CA* by both PCR methods had stage IV disease and a PFS significantly poorer than average for the cohort. Another patient with high cfDNA levels and amplification of all genes had a high-frequency mutation in *TP53*, which is suggestive of aneuploidy.

Of particular interest, correlation of SCNA/NGS data with survival outcomes revealed that patients with SCNA, mutation or both types of alteration had significantly worse OS and PFS than in patients without any detectable alterations. This was more significant for SCNA-only than mutation-only patients, and was most significant when patients harboured both types of alteration. This clearly demonstrates the value of SCNA analysis (without or in combination with NGS) in the GALAXY-1 cohort in predicting survival outcomes, and is the first time that SCNA analysis in cfDNA has shown such utility. Moreover, PCR-based SCNA panels such as the panel developed in this thesis may find clinical utility, owing to the simple workflow and low cost-per-sample (relative to NGS).

### 5.4.3     Sensitivity of methods used to evaluate SCNA in cfDNA

It is important to consider the sensitivity of SCNA detection methods in discriminating fine changes in copy number, since SCNA in blood is brought closer to the germline copy number by the dilution effect of healthy cfDNA. For example, when a tumour-specific SCNA with a CNR of 5.0 represents 5% of total cfDNA, the CNR is reduced to 1.2 (Whale *et al.*, 2012). As previously discussed, the reliance of qPCR on equal efficiency between target and reference assays may not permit an accurate quantitation of copy number, and low amplification thresholds (such as CNR, 1.2) may fail to discriminate between true and error-induced copy number gains. In contrast, several studies have shown the ability of ddPCR to distinguish between amplified and non-amplified tumours with respect to a particular locus, using a low threshold for amplification in cfDNA. Gevensleben *et al.* used a CNR threshold of 1.25 to indicate *HER2* amplification in cfDNA of patients with metastatic breast cancer, which identified amplification in 64% of confirmed *HER2*-positive patients and no amplification in 94% of *HER2*-negative patients (Gevensleben *et al.*, 2013). Kinugasa *et al.* evaluated *HER2* amplification in serum and FFPE tissue in patients with gastric cancer by ddPCR. Using a CNR threshold of 1.20, the authors achieved a concordance rate of 63% with respect to amplification status (Kinugasa *et al.*, 2015). Recent targeted and whole-exome NGS studies have reported a SCNA detection limit of 10% ctDNA (Page *et al.*, 2016 [in press]) (Belic *et al.*, 2015). This figure is in agreement with previous NGS analyses (Heitzer *et al.*, 2013; Murtaza *et al.*, 2013; Mohan *et al.*, 2014). A recent study used a novel NGS strategy to target SNPs in five chromosome arms subject to frequent SCNA in breast cancer. The authors demonstrated sensitivity down to an average allelic imbalance of 0.5% in cfDNA of patients with stage II disease, which is at least five times more sensitive than whole-genome sequencing (Kirkizlar *et al.*, 2015). Therefore, with ctDNA comprising between 0.01% and 90% of cfDNA, it is possible that SCNA in some cancers may not be detectable with even the most sensitive detection methods (Schwarzenbach *et al.*, 2008).

## 5.5       Conclusions

The work presented in this chapter has demonstrated the ability of a PCR-based SCNA gene panel to identify alterations in cfDNA of a subset of patients with locally advanced/advanced NSCLC, the majority of which did not have mutations detected by NGS. Moreover, correlation of SCNA and mutation data with survival characteristics in the GALAXY-1 cohort revealed significantly reduced OS and PFS in patients with SCNA and/or mutations, relative to patients with no detectable alterations in cfDNA. Interestingly, SCNA status alone identified patients with significantly poorer OS, and this is the first time SCNA in cfDNA has been shown to correlate with survival outcomes in any cancer type.

## 5.6       Future direction

Since most patients with SCNA did not have detectable mutations, SCNA analysis identified an additional subset of patients despite interrogating only five commonly amplified genes. Future work may involve expanding the gene panel to include more loci with the aim of detecting ctDNA in a greater proportion of patients, in addition to evaluating SCNA in patients with earlier stage disease. The significant correlation between SCNA status and OS in the GALAXY-1 cohort may prompt further studies assessing the utility of SCNA in cfDNA as a predictor of survival in cancer.

# Chapter 6

# General Discussion

**6.1 Analysis of cfDNA in GEM models of NSCLC**

Around two-thirds of patients with non-small cell lung cancer (NSCLC) present with advanced disease, which is refractory to curative treatment (Folkert and Timmerman, 2015). Therefore, early detection strategies for NSCLC remain a major unmet need for improving clinical outcomes (Birring and Peake, 2005). Owing to the largely asymptomatic nature of early NSCLC, research into early detection of the disease is hampered by the inability to identify individuals with either early disease or individuals harbouring preneoplastic lesions that may progress to malignancy. One research approach is to use genetically engineered mouse (GEM) models of cancer, in which activation of one or more oncogenes is under precise spatiotemporal control. As suggested by Kelly-Spratt *et al.* (2008), clinically-relevant GEM models may be ideal platforms for biomarker evaluation (including cf/ctDNA), since they can recapitulate many features of the cognate human diseases while eliminating sources of confounding genetic and environmental variability (Kelly-Spratt *et al.*, 2008). Most advantageous to early detection research is the ability to monitor *de novo* tumour development from initiation and obtain blood (e.g. for isolation of cfDNA) at desired time points for comparison with disease burden.

There are currently no published studies that report use of GEM models for cfDNA analysis. In this thesis, $Braf^{V600E}$- and $Kras^{G12D}$-driven GEM models of early lung adenocarcinoma development were used for longitudinal evaluation of cfDNA in comparison with disease burden. These models are controlled spatiotemporally by Cre-Lox recombination, and develop preneoplastic lesions recapitulating those observed in humans (Dankort *et al.*, 2007; Johnson *et al.*, 2001). Proof-of-concept work using conventional end-point PCR demonstrated the presence of lesion-derived cfDNA through detection of the recombined conditional allele (single, recombined LoxP sequence) in both models 12 weeks post-adenoviral Cre administration. Moreover, $Kras^{G12D}$ mice had significantly elevated total cfDNA levels relative to wild-type mice by qPCR. Of note, these observations were made in mice with preneoplastic lung lesions but no evidence of malignant disease, suggesting that it may be possible to detect mutations in

preinvasive human malignancies through cfDNA. However, several aspects of this proof-of-concept study were different to cfDNA analysis in humans. First, lesion burden (up to 90% of lung volume) was significantly higher than observed in humans. Second, cfDNA was isolated from blood collected by cardiac puncture, a technique that can yield up to two-thirds of total blood volume (McClure, 1999). To improve clinical relevance, further studies aimed to quantitatively assess cfDNA in $Kras^{G12D}$ mice with lower tumour burden from small volumes of blood collected at defined time points, and to quantify tumour burden by computed tomography (CT) imaging.

To permit quantitative assessment of Cre-Lox recombination in cfDNA, PCR assays were designed to target unrecombined and recombined alleles of $Kras^{G12D}$. Weis *et al.* (2010) published a qPCR assay to detect gene deletion in a Cre-Lox regulated GEM model by targeting the recombined LoxP sequence (Weis *et al.*, 2010). In this thesis, an assay to detect the unrecombined allele of $Kras^{G12D}$ was successfully developed based on the Weis *et al.* assay. However, this methodology could not be used to detect the recombined allele, because the recombined LoxP sequence was predicted to form a thermostable stem-loop structure that prevented PCR amplification (Zuker, 2003). Amplification could not be achieved using high-$T_m$ primers and/or hydrolysis probes, but was achieved using a locked nucleic acid (LNA)-modified probe targeted to LoxP sequences on both strands, promoting destabilisation of both stem-loops. This strategy was based on a published method by Esposito *et al.* (2003) for improving sequencing through *att*L recombination sites, which are inverted repeat sequences that form stem-loops (Esposito *et al.*, 2003). Using droplet digital PCR (ddPCR), the recombined $Kras^{G12D}$ detection assay demonstrated a limit of detection of 0.1% in a background of 3.7 ng unrecombined allele, and was therefore deemed suitable for analysis of cfDNA. The PCR assays developed in this thesis for quantitative detection of Cre-Lox recombination in cfDNA may have broad applicability to other GEM models of cancer, including models used in ongoing cfDNA projects at the University of Leicester.

One potential confounding factor of cfDNA analysis using the PCR assays previously described is that mutations may occur within the LoxP sequence. Therefore, the recombined LoxP sequence and surrounding genomic region were cloned from tumour-bearing lung tissue of a *Kras*$^{G12D}$ mouse, and analysis of 40 clones revealed a high frequency of point mutations, both within the LoxP sequence and in flanking intronic regions of *Kras*$^{G12D}$. These mutations were not observed in the corresponding intronic regions of *Kras*$^{WT}$ from the same tissue sample. Previous studies have demonstrated that point mutations in the 13 bp Cre-binding sequences of LoxP inhibit recombination, unless two unrecombined LoxP sequences have identical mutations (Thomson *et al.*, 2003; Chatterjee *et al.*, 2004). Taken together, it is unlikely that the mutations observed in recombined *Kras*$^{G12D}$ were already present in the unrecombined LoxP sequences prior to Cre-mediated recombination.

To permit correlation of cfDNA with disease stage, tumour burden was assessed by CT imaging at defined time points from activation of *Kras*$^{G12D}$ to end-point. The *Kras*$^{G12D}$ model has been well characterised by CT imaging previously (Haines *et al.*, 2009; Lalwani *et al.*, 2013; Cavanaugh *et al.*, 2004; Kirsch *et al.*, 2010). In agreement with previous reports, volumetric analysis of individual lung lesions was not possible due to their small size, poorly circumscribed borders and the confounding effect of airway and vasculature structures (Haines *et al.*, 2009; Lalwani *et al.*, 2013). Instead, a method reported by Haines *et al.* (2009) was used to calculate tumour burden, in which the combined volume of tumour and vasculature was expressed as a percentage of total lung volume (Haines *et al.*, 2009). Using this method, tumour burden was found to increase over time in *Kras*$^{G12D}$ but not wild-type mice, thus serving as a means to follow progression of tumours and correlate with quantitative changes in the cfDNA profile.

Unfortunately, serial cfDNA extracts became contaminated with plasmid DNA (containing recombined *Kras*$^{G12D}$), most likely via a contaminated pipette, which precluded correlation analysis between tumour burden and levels of recombined *Kras*$^{G12D}$ in cfDNA over time. Moreover, total levels of cfDNA in longitudinal blood samples of *Kras*$^{G12D}$ mice varied. There were some issues with sample collection,

including obvious haemolysis revealed by a red-tinted plasma; these samples generally yielded very high levels of (cf)DNA due to contamination with genomic DNA. Future studies should seek to address the methodological challenges of handling small blood volumes that may lead to such erroneous results. Despite removing outlier samples, no trends in cfDNA levels were seen over time in *Kras*$^{G12D}$ mice or wild-type mice.

## 6.2      Significance of SCNA in cfDNA of patients with NSCLC

Several oncogenes are subject to frequent amplification in NSCLC, and some amplifications are associated with clinical outcome. For example, amplifications of *EGFR* and *MET* confer poor prognosis, and *EGFR* amplification also confers sensitivity to tyrosine kinase inhibitors (Cappuzzo *et al.*, 2005; Hirsch *et al.*, 2005; Beau-Faller *et al.*, 2008; Zhu *et al.*, 2008; Hirsch *et al.*, 2006). Moreover, *MET* amplification is a known mechanism of resistance to such drugs (Go *et al.*, 2010; Cappuzzo *et al.*, 2009b; Okuda *et al.*, 2008). Therefore, detection of these alterations in cfDNA may enable non-invasive disease management.

Previous studies from this and other groups have demonstrated amplification at discrete genomic loci in cfDNA of patients with breast and lung cancers using qPCR and ddPCR (Page *et al.*, 2011; Gevensleben *et al.*, 2013; Beck *et al.*, 2013) (Page *et al.*, 2016). In this study, qPCR was used to evaluate somatic copy number alteration (SCNA) in six genes that are commonly amplified in NSCLCs (*EGFR*, *FGFR1*, *HER2*, *MET*, *PIK3CA* and *SOX2*) in cfDNA from 183 patients with stage IIIA – IV NSCLC. ddPCR was used as a comparator method in samples that showed SCNA by qPCR. Fifteen percent of patients showed amplification in at least one gene. Interestingly, a third of SCNA-positive samples showed coamplification of two or more genes, most commonly *EGFR* and *MET*; previous studies have reported coamplification of these genes in 20 – 30% of *MET*-amplified NSCLC (Cappuzzo *et al.*, 2009a; Beau-Faller *et al.*, 2008). Other coamplified loci included *PIK3CA* and *MET*, and *FGFR1* and *HER2*, which have not been reported previously.

Consistent with previous reports, there was a general downward trend in copy number measurements by ddPCR relative to qPCR across all genes (Whale *et al.*, 2012; Devonshire *et al.*, 2014; Bharuthram *et al.*, 2014). A limitation of qPCR for copy number analysis was evident by observing the spread of copy number values across the cohort for *PIK3CA.* There was a 4% difference in PCR efficiency between assays for *PIK3CA* and *RPPH1* and, as a result, copy number values for *PIK3CA* were more scattered than for any other gene. Moreover, *PIK3CA* showed the greatest discordance between copy numbers calculated by qPCR and ddPCR. In line with this observation, Ramakers *et al.* (2003) reported a bias in $\Delta C_T$ of up to 400% with a 4% difference in efficiency between target and reference (Ramakers *et al.*, 2003). Therefore qPCR data for *PIK3CA* must be excluded.

Combined analysis of SCNA and mutations (identified through Ion Torrent™ NGS in 163 patients) revealed that only 4 patient samples had both SCNA and mutations in cfDNA, thus demonstrating the added value of combined SNV and SCNA analysis, with each analysis detecting ctDNA in a distinct group of patients. Significant correlations were observed between SCNA and/or mutation status in cfDNA and survival outcomes. Patients with SCNA in cfDNA had significantly worse overall survival than those patients with no SCNA detected. The same trend was seen, for patients with mutations but no SCNA detected in cfDNA. Overall, combining SCNA and mutation data showed that patients with any detectable alteration in cfDNA had poorer progression-free survival (PFS) and overall survival (OS) than patients with neither alteration.

Numerous studies have investigated the predictive capacity of cfDNA biomarkers in patients with NSCLC (Catarino *et al.*, 2012; Li *et al.*, 2016; Nygaard *et al.*, 2014; Sirera *et al.*, 2011; Camps *et al.*, 2006; Gautschi *et al.*, 2004; Vinayanuwattikun *et al.*, 2013). A recent meta-analysis confirmed that elevated cfDNA levels are a predictor of poor PFS (HR, 1.32; 95% CI, 0.66 – 2.56) and OS (HR, 1.64; 95% CI, 0.49 – 2.25), presumably because cfDNA levels are correlated with tumour burden (Ai *et al.*, 2016). However, this review did not find a significant predictive role of tumour-specific alterations (*EGFR* and *KRAS* mutations) in cfDNA. To my

knowledge, there are no publications to date that describe the use of SCNAs in cfDNA as predictive biomarkers in patients with NSCLC. Data described in this thesis suggests the utility of specific SCNAs as predictive biomarkers of cancer through liquid biopsy, and warrants validation in larger cohorts of patients. If validated, this raises the prospect of using simple, low-cost PCR-based gene panels in the clinical setting.

## 6.3    Conclusion

In conclusion, data described in this thesis has demonstrated proof-of-concept that tumour-derived alterations can be detected in cfDNA of $Braf^{V600E}$ and $Kras^{G12D}$-driven GEM models of NSCLC in mice with preneoplastic lesions alone. Moreover, PCR strategies have been developed for quantitative detection of Cre-Lox recombination in cfDNA of $Kras^{G12D}$ mice, in addition to measurement of total cfDNA levels, permitting longitudinal analysis of cfDNA in comparison with tumour burden by CT imaging. In an unrelated study, SCNA analysis was able to detect ctDNA in a group of patients with advanced NSCLC that did not show evidence of hotspot mutations in 50 cancer genes in ctDNA by NGS. Patients with detectable ctDNA had significantly poorer survival outcomes than patients with no detected SCNA in cfDNA. This demonstrates, for the first time, the potential of SCNAs as predictive biomarkers of NSCLC.

## 6.4        Future direction

Based on the observation that tumour-derived cfDNA was detected in GEM models with lung preneoplasia, similar investigation is warranted in GEM models of other cancer types, such as pancreatic cancer, where early detection strategies are urgently sought. Moreover, since $Kras^{G12D}$ is a common driver of pancreatic ductal adenocarcinoma, and pancreatic expression of $Kras^{G12D}$ leads to formation of preneoplastic lesions in mice, the PCR assays developed for quantitative detection of Cre-Lox recombination in this thesis are directly transferable to analysis of cfDNA in such models (Hingorani *et al.*, 2003; Carriere *et al.*, 2007; Gidekel Friedlander *et al.*, 2009). Also, the PCR assays may be used to investigate cfDNA in GEM models of lung adenocarcinoma driven by $Kras^{G12D}$ and concomitant loss of tumour suppressors, such as *Trp53*, *Ink4a/Arf* and *Lkb1* (Jackson *et al.*, 2005; Ji *et al.*, 2007; Fisher *et al.*, 2001). Sequence variation within and around the recombined LoxP sequence is worthy of further investigation in Cre-Lox regulated alleles of other GEM models, to determine whether the variation observed in this thesis was specific to the $Kras^{G12D}$ model, or whether such mutations arise through Cre-mediated recombination. A major challenge in analysis of mouse cfDNA in this thesis was the collection and handling of small volumes of blood (< 40 µL), with samples frequently displaying signs of haemolysis, and contamination of plasma with white blood cells. Therefore, future studies should address the methodology relating to blood collection and handling, in order to permit a more standardised quantitative measurement of cfDNA levels over time.

Since most patients with SCNA in cfDNA did not have detectable mutations, SCNA analysis identified an additional subset of patients with NSCLC with detectable ctDNA, despite interrogating only five commonly amplified genes. Future work could involve expanding the gene panel to include other loci, in addition to evaluating SCNA in patients with earlier stage disease. The significant correlation between SCNA status and overall survival in the NSCLC patient cohort may prompt further studies with large patient cohorts, assessing the utility of SCNA in cfDNA as a predictor of survival in cancer.

# Appendix

| ID | Tx arm | Sex | Stg | Smoking status | *KRAS* status | Death | OS (days) | PFS (days) | ID | Tx arm | Sex | Stg | Smoking status | *KRAS* status | Death | OS (days) | PFS (days) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P001 | D | F | IV | Never | None | NO | 621 | 549 | P073 | D/G | M | IV | Past | None | YES | 244 | 119 |
| P002 | D/G | F | IV | Never | None | YES | 168 | 88 | P074 | D/G | M | IV | Current | None | NO | 527 | 337 |
| P003 | D | M | IV | Current | None | YES | 72 | 72 | P075 | D | M | IV | Current | G12C | NO | 509 | 266 |
| P004 | D | F | IIIB | Past | G12V | NO | 608 | 169 | P076 | D | M | IV | Past | None | NO | 56 | 21 |
| P005 | D/G | F | IIIB | Past | G13D | YES | 350 | 88 | P077 | D/G | F | IV | Never | None | YES | 228 | 127 |
| P006 | D | M | IV | Current | G12V | YES | 33 | 33 | P078 | D/G | M | IV | Current | None | NO | 526 | 510 |
| P007 | D/G | M | IV | Past | None | YES | 60 | 60 | P079 | D | M | IIIB | Past | None | YES | 524 | 51 |
| P008 | D | F | IV | Past | None | NO | 711 | 701 | P080 | D | F | IIIB | Never | None | YES | 94 | 31 |
| P009 | D/G | M | IV | Current | None | YES | 8 | 8 | P081 | D/G | F | IV | Past | None | YES | 208 | 75 |
| P010 | D | F | IV | Current | G12V | YES | 195 | 42 | P082 | D/G | F | IV | Never | None | YES | 246 | 169 |
| P011 | D | M | IV | Past | G12C | YES | 121 | 44 | P083 | D/G | M | IV | Past | None | NO | 478 | 171 |
| P012 | D/G | F | IV | Never | None | NO | 542 | 338 | P084 | D/G | M | IV | Past | G12C | NO | 599 | 420 |
| P013 | D | F | IV | Never | None | YES | 275 | 209 | P085 | D/G | M | IV | Past | None | YES | 160 | 160 |
| P014 | D | M | IV | Past | None | NO | 534 | 421 | P086 | D | F | IV | Never | None | YES | 120 | 112 |
| P015 | D | M | IV | Current | None | YES | 237 | 209 | P087 | D | M | IV | Past | None | YES | 95 | 1 |
| P016 | D/G | F | IV | Current | None | NO | 531 | 126 | P088 | D | M | IV | Past | G12A | YES | 118 | 118 |
| P017 | D | F | IV | Past | None | YES | 244 | 244 | P089 | D | M | IV | Past | None | YES | 309 | 128 |
| P018 | D | M | IIIB | Past | None | YES | 59 | 19 | P090 | D | M | IV | Past | G12C | YES | 363 | 209 |
| P019 | D/G | M | IV | Past | None | NO | 545 | 339 | P091 | D | M | IV | Current | G12C | YES | 87 | 82 |
| P020 | D | M | IV | Current | None | YES | 86 | 43 | P092 | D/G | F | IV | Past | None | NO | 664 | 323 |
| P021 | D | F | IV | Past | G12C | YES | 177 | 87 | P093 | D/G | M | IV | Never | G12D | YES | 231 | 170 |
| P022 | D | M | IV | Current | None | YES | 71 | 43 | P094 | D/G | M | IV | Past | G12C | YES | 68 | 22 |
| P023 | D/G | F | IV | Past | None | YES | 358 | 88 | P095 | D/G | M | IV | Past | G12V | NO | 617 | 118 |
| P024 | D | M | IV | Current | G12C | YES | 135 | 41 | P096 | D/G | F | IV | Past | G12V | NO | 583 | 525 |
| P025 | D | F | IV | Past | None | YES | 154 | 42 | P097 | D/G | F | IV | Never | None | YES | 34 | 17 |
| P026 | D | M | IV | Past | None | YES | 366 | 216 | P098 | D | F | IV | Current | None | YES | 87 | 42 |
| P027 | D/G | M | IV | Past | None | YES | 136 | 44 | P099 | D | M | IV | Past | G12V | NO | 364 | 211 |
| P028 | D | F | IV | Past | None | NO | 520 | 520 | P100 | D/G | M | IV | Current | G12D | YES | 10 | 10 |
| P029 | D/G | F | IV | Never | G12V | NO | 509 | 169 | P101 | D/G | M | IV | Past | G12C | YES | 335 | 85 |
| P030 | D | F | IV | Never | None | YES | 29 | 19 | P102 | D/G | M | IV | Past | None | NO | 360 | 85 |
| P031 | D/G | M | IV | Past | None | YES | 185 | 138 | P103 | D | M | IV | Current | None | YES | 6 | 6 |
| P032 | D/G | F | IV | Past | None | NO | 534 | 173 | P104 | D | F | IV | Never | None | YES | 317 | 85 |
| P033 | D | M | IV | Past | G12C | YES | 30 | 30 | P105 | D/G | M | IV | Never | G12V | YES | 21 | 21 |
| P034 | D | M | IV | Never | None | NO | 477 | 138 | P106 | D/G | M | IV | Past | G12V | YES | 15 | 15 |
| P035 | D | M | IV | Current | None | YES | 406 | 253 | P107 | D/G | F | IV | Never | None | NO | 355 | 341 |
| P036 | D/G | F | IV | Never | None | NO | 196 | 86 | P108 | D/G | M | IV | Past | None | YES | 456 | 213 |
| P037 | D/G | M | IV | Current | None | YES | 48 | 48 | P109 | D | F | IV | Past | None | YES | 239 | 22 |
| P038 | D | M | IV | Current | G12C | NO | 461 | 131 | P110 | D | M | IV | Current | G12V | YES | 72 | 69 |
| P039 | D | M | IV | Past | None | NO | 429 | 85 | P111 | D/G | F | IV | Never | None | YES | 276 | 179 |
| P040 | D | F | IV | Current | None | YES | 131 | 21 | P112 | D/G | F | IV | Past | None | NO | 540 | 465 |
| P041 | D | M | IV | Never | None | YES | 206 | 82 | P113 | D | M | IV | Past | None | YES | 57 | 57 |
| P042 | D | M | IV | Current | None | NO | 3 | 1 | P114 | D/G | M | IV | Current | None | NO | 48 | 20 |
| P043 | D | F | IV | Current | G12D | YES | 278 | 1 | P115 | D/G | M | IV | Past | G12V | YES | 107 | 43 |
| P044 | D | F | IV | Never | None | NO | 351 | 215 | P116 | D/G | F | IV | Past | None | NO | 530 | 380 |
| P045 | D | M | IV | Past | G12V | YES | 65 | 11 | P117 | D | M | IV | Past | None | YES | 107 | 23 |
| P046 | D/G | M | IV | Past | None | YES | 24 | 1 | P118 | D | F | IV | Past | None | YES | 137 | 87 |
| P047 | D | F | IV | Past | G12V | YES | 130 | 90 | P119 | D/G | F | IV | Past | None | NO | 447 | 255 |
| P048 | D | M | IV | Past | None | YES | 34 | 34 | P120 | D/G | M | IV | Never | G12D | YES | 17 | 17 |
| P049 | D/G | F | IV | Past | G12C | YES | 294 | 89 | P121 | D/G | M | IV | Past | None | NO | 383 | 340 |
| P050 | D/G | F | IV | Never | None | NO | 328 | 301 | P122 | D | M | IV | Never | G12D | YES | 37 | 35 |
| P051 | D | F | IV | Current | G12C | YES | 158 | 96 | P123 | D/G | F | IV | Current | G12D | NO | 551 | 85 |
| P052 | D/G | F | IV | Never | G12C | YES | 326 | 211 | P124 | D/G | M | IV | Current | G12V | YES | 130 | 130 |
| P053 | D | M | IV | Past | None | YES | 125 | 22 | P125 | D/G | M | IIIB | Past | None | YES | 117 | 102 |
| P054 | D | M | IV | Past | G12D | YES | 350 | 172 | P126 | D/G | F | IV | Past | None | NO | 503 | 208 |
| P055 | D/G | M | IV | Current | G12C | YES | 322 | 126 | P127 | D | F | IV | Never | None | NO | 427 | 258 |
| P056 | D | F | IV | Never | None | YES | 69 | 69 | P128 | D | F | IV | Never | None | YES | 309 | 37 |
| P057 | D | M | IV | Past | None | YES | 174 | 127 | P129 | D/G | M | IV | Current | G12S | YES | 78 | 23 |
| P058 | D | M | IV | Current | None | YES | 404 | 211 | P130 | D | M | IV | Past | None | YES | 340 | 172 |
| P059 | D/G | F | IV | Never | None | YES | 440 | 169 | P131 | D | M | IV | Past | G12C | YES | 147 | 83 |
| P060 | D | M | IV | Current | None | NO | 540 | 128 | P132 | D/G | M | IV | Never | None | YES | 120 | 65 |
| P061 | D/G | F | IV | Never | None | YES | 227 | 160 | P133 | D/G | M | IV | Current | G12C | YES | 102 | 44 |
| P062 | D | M | IV | Past | None | NO | 657 | 211 | P134 | D | F | IV | Never | None | YES | 364 | 85 |
| P063 | D | M | IV | Current | None | YES | 155 | 83 | P135 | D/G | M | IV | Past | None | YES | 441 | 211 |
| P064 | D | M | IV | Never | None | NO | 674 | 468 | P136 | D/G | F | IV | Past | G12C | YES | 367 | 89 |
| P065 | D | F | IV | Past | G12C | YES | 9 | 9 | P137 | D | F | IV | Current | G12C | NO | 211 | 167 |
| P066 | D | F | IV | Current | None | YES | 254 | 125 | P138 | D | M | IV | Past | None | YES | 72 | 19 |
| P067 | D/G | M | IV | Past | G12V | YES | 53 | 22 | P139 | D/G | M | IV | Past | None | YES | 58 | 37 |
| P068 | D | M | IV | Past | None | NO | 605 | 80 | P140 | D/G | M | IIIB | Past | None | NO | 554 | 505 |
| P069 | D | M | IV | Past | None | YES | 152 | 21 | P141 | D | M | IV | Past | None | YES | 100 | 57 |
| P070 | D | M | IV | Never | None | NO | 632 | 632 | P142 | D/G | M | IIIB | Current | None | YES | 310 | 22 |
| P071 | D | M | IV | Current | None | YES | 353 | 83 | P143 | D | M | IV | Past | None | YES | 442 | 251 |
| P072 | D/G | M | IV | Past | G12C | YES | 43 | 20 | | | | | | | | | |

**Supplementary Table S1.1 – Detailed clinicopathological characteristics of GALAXY-1 patients.** D, docetaxel; D/G, docetaxel and ganetespib; OS, overall survival; PFS, progression-free survival; Tx arm, treatment arm.

# **Bibliography**

Ai, B., Liu, H., Huang, Y., Peng, P., 2016. Circulating cell-free DNA as a prognostic and predictive biomarker in non-small cell lung cancer. *Oncotarget.*

Aldhous, M.C., Abu Bakar, S., Prescott, N.J., Palla, R., Soo, K., Mansfield, J.C., Mathew, C.G., Satsangi, J., Armour, J.A., 2010. Measurement methods and accuracy in copy number variation: failure to replicate associations of beta-defensin copy number with Crohn's disease. *Human Molecular Genetics.* **19,** 4930-4938.

Amann, J., Kalyankrishna, S., Massion, P.P., Ohm, J.E., Girard, L., Shigematsu, H., Peyton, M., Juroske, D., Huang, Y., Stuart Salmon, J., Kim, Y.H., Pollack, J.R., Yanagisawa, K., Gazdar, A., Minna, J.D., Kurie, J.M., Carbone, D.P., 2005. Aberrant epidermal growth factor receptor signaling and enhanced sensitivity to EGFR inhibitors in lung cancer. *Cancer Research.* **65,** 226-235.

An, S.J., Chen, Z.H., Su, J., Zhang, X.C., Zhong, W.Z., Yang, J.J., Zhou, Q., Yang, X.N., Huang, L., Guan, J.L., Nie, Q., Yan, H.H., Mok, T.S., Wu, Y.L., 2012. Identification of enriched driver gene alterations in subgroups of non-small cell lung cancer patients based on histology and smoking status. *PloS One.* **7,** e40109.

Andriani, F., Conte, D., Mastrangelo, T., Leon, M., Ratcliffe, C., Roz, L., Pelosi, G., Goldstraw, P., Sozzi, G., Pastorino, U., 2004. Detecting lung cancer in plasma with the use of multiple genetic markers. *International Journal of Cancer.Journal International Du Cancer.* **108,** 91-96.

Antczak, A., Migdalska-Sek, M., Pastuszak-Lewandoska, D., Czarnecka, K., Nawrot, E., Domanska, D., Kordiak, J., Gorski, P., Brzezianska, E., 2013. Significant frequency of allelic imbalance in 3p region covering RARbeta and MLH1 loci seems to be essential in molecular non-small cell lung cancer diagnosis. *Medical Oncology (Northwood, London, England).* **30,** 532-013-0532-9. Epub 2013 Mar 17.

Aranda, M., Kanellopoulou, C., Christ, N., Peitz, M., Rajewsky, K., Droge, P., 2001. Altered directionality in the Cre-LoxP site-specific recombination pathway. *Journal of Molecular Biology.* **311,** 453-459.

Araujo, L.H., Timmers, C., Bell, E.H., Shilo, K., Lammers, P.E., Zhao, W., Natarajan, T.G., Miller, C.J., Zhang, J., Yilmaz, A.S., Liu, T., Coombes, K., Amann, J., Carbone, D.P., 2015. Genomic Characterization of Non-Small-Cell Lung Cancer in African Americans by Targeted Massively Parallel Sequencing. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **33,** 1966-1973.

Ardizzoni, A., Hansen, H., Dombernowsky, P., Gamucci, T., Kaplan, S., Postmus, P., Giaccone, G., Schaefer, B., Wanders, J., Verweij, J., 1997. Topotecan, a new active drug in the second-line treatment of small-cell lung

cancer: a phase II study in patients with refractory and sensitive disease. The European Organization for Research and Treatment of Cancer Early Clinical Studies Group and New Drug Development Office, and the Lung Cancer Cooperative Group. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **15,** 2090-2096.

Armour, J.A., Palla, R., Zeeuwen, P.L., den Heijer, M., Schalkwijk, J., Hollox, E.J., 2007. Accurate, high-throughput typing of copy number variation using paralogue ratios from dispersed repeats. *Nucleic Acids Research.* **35,** e19.

Arriola, E., Canadas, I., Arumi, M., Rojo, F., Rovira, A., Albanell, J., 2008. Genetic changes in small cell lung carcinoma. *Clinical and Translational Oncology : Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico.* **10,** 189-197.

Ashton, J.R., Clark, D.P., Moding, E.J., Ghaghada, K., Kirsch, D.G., West, J.L., Badea, C.T., 2014. Dual-energy micro-CT functional imaging of primary lung cancer in mice using gold and iodine nanoparticle contrast agents: a validation study. *PloS One.* **9,** e88129.

Badea, C.T., Athreya, K.K., Espinosa, G., Clark, D., Ghafoori, A.P., Li, Y., Kirsch, D.G., Johnson, G.A., Annapragada, A., Ghaghada, K.B., 2012. Computed tomography imaging of primary lung cancer in mice using a liposomal-iodinated contrast agent. *PloS One.* **7,** e34496.

Bass, A.J., Watanabe, H., Mermel, C.H., Yu, S., Perner, S., Verhaak, R.G., Kim, S.Y., Wardwell, L., Tamayo, P., Gat-Viks, I., Ramos, A.H., Woo, M.S., Weir, B.A., Getz, G., Beroukhim, R., O'Kelly, M., Dutt, A., Rozenblatt-Rosen, O., Dziunycz, P., Komisarof, J., Chirieac, L.R., Lafargue, C.J., Scheble, V., Wilbertz, T., Ma, C., Rao, S., Nakagawa, H., Stairs, D.B., Lin, L., Giordano, T.J., Wagner, P., Minna, J.D., Gazdar, A.F., Zhu, C.Q., Brose, M.S., Cecconello, I., Jr, U.R., Marie, S.K., Dahl, O., Shivdasani, R.A., Tsao, M.S., Rubin, M.A., Wong, K.K., Regev, A., Hahn, W.C., Beer, D.G., Rustgi, A.K., Meyerson, M., 2009. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nature Genetics.* **41,** 1238-1242.

Basu, D. and Herlyn, M., 2009. Defining microenvironments within mouse models that enhance tumor aggressiveness. *Cancer Biology and Therapy.* **8,** 380-381.

Bean, J., Brennan, C., Shih, J.Y., Riely, G., Viale, A., Wang, L., Chitale, D., Motoi, N., Szoke, J., Broderick, S., Balak, M., Chang, W.C., Yu, C.J., Gazdar, A., Pass, H., Rusch, V., Gerald, W., Huang, S.F., Yang, P.C., Miller, V., Ladanyi, M., Yang, C.H., Pao, W., 2007. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proceedings of the National Academy of Sciences of the United States of America.* **104,** 20932-20937.

Beau-Faller, M., Gaub, M.P., Schneider, A., Ducrocq, X., Massard, G., Gasser, B., Chenard, M.P., Kessler, R., Anker, P., Stroun, M., Weitzenblum, E., Pauli, G., Wihlm, J.M., Quoix, E., Oudet, P., 2003. Plasma DNA microsatellite panel as sensitive and tumor-specific marker in lung cancer patients. *International Journal of Cancer.Journal International Du Cancer.* **105,** 361-370.

Beau-Faller, M., Ruppert, A.M., Voegeli, A.C., Neuville, A., Meyer, N., Guerin, E., Legrain, M., Mennecier, B., Wihlm, J.M., Massard, G., Quoix, E., Oudet, P., Gaub, M.P., 2008. MET gene copy number in non-small cell lung cancer: molecular analysis in a targeted tyrosine kinase inhibitor naive cohort. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **3,** 331-339.

Beck, J., Hennecke, S., Bornemann-Kolatzki, K., Urnovitz, H.B., Neumann, S., Strobel, P., Kaup, F.J., Brenig, B., Schutz, E., 2013. Genome aberrations in canine mammary carcinomas and their detection in cell-free plasma DNA. *PloS One.* **8,** e75485.

Bednarz, N., Blaut, K., Sworczak, K., Oseka, T., Bielawski, K.P., 2009. The profile of ErbB/Her family genes copy number assessed by real-time PCR in parathyroid adenoma and hyperplasia associated with sporadic primary hyperparathyroidism. *Acta Biochimica Polonica.* **56,** 83-88.

Belic, J., Koch, M., Ulz, P., Auer, M., Gerhalter, T., Mohan, S., Fischereder, K., Petru, E., Bauernhofer, T., Geigl, J.B., Speicher, M.R., Heitzer, E., 2015. Rapid Identification of Plasma DNA Samples with Increased ctDNA Levels by a Modified FAST-SeqS Approach. *Clinical Chemistry.* **61,** 838-849.

Benjamini, Y., Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological).* **57**, 289-300.

Bergethon, K., Shaw, A.T., Ou, S.H., Katayama, R., Lovly, C.M., McDonald, N.T., Massion, P.P., Siwak-Tapp, C., Gonzalez, A., Fang, R., Mark, E.J., Batten, J.M., Chen, H., Wilner, K.D., Kwak, E.L., Clark, J.W., Carbone, D.P., Ji, H., Engelman, J.A., Mino-Kenudson, M., Pao, W., Iafrate, A.J., 2012. ROS1 rearrangements define a unique molecular class of lung cancers. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **30,** 863-870.

Bharuthram, A., Paximadis, M., Picton, A.C., Tiemessen, C.T., 2014. Comparison of a quantitative Real-Time PCR assay and droplet digital PCR for copy number analysis of the CCL4L genes. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases.* **25,** 28-35.

Bidard, F.C., Madic, J., Mariani, P., Piperno-Neumann, S., Rampanou, A., Servois, V., Cassoux, N., Desjardins, L., Milder, M., Vaucher, I., Pierga, J.Y., Lebofsky, R., Stern, M.H., Lantz, O., 2014. Detection rate and prognostic value of circulating tumor cells and circulating tumor DNA in metastatic uveal melanoma. *International Journal of Cancer.Journal International Du Cancer.* **134,** 1207-1213.

Birring, S.S. and Peake, M.D., 2005. Symptoms and the early diagnosis of lung cancer. *Thorax.* **60,** 268-269.

Boers, J.E., Ambergen, A.W., Thunnissen, F.B., 1998. Number and proliferation of basal and parabasal cells in normal human airway epithelium. *American Journal of Respiratory and Critical Care Medicine.* **157,** 2000-2006.

Borczuk, A.C., Qian, F., Kazeros, A., Eleazar, J., Assaad, A., Sonett, J.R., Ginsburg, M., Gorenstein, L., Powell, C.A., 2009. Invasive size is an independent predictor of survival in pulmonary adenocarcinoma. *The American Journal of Surgical Pathology.* **33,** 462-469.

Brambilla, E., Gazzeri, S., Lantuejoul, S., Coll, J.L., Moro, D., Negoescu, A., Brambilla, C., 1998. p53 mutant immunophenotype and deregulation of p53 transcription pathway (Bcl2, Bax, and Waf1) in precursor bronchial lesions of lung cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **4,** 1609-1618.

Brcic, L., Sherer, C.K., Shuai, Y., Hornick, J.L., Chirieac, L.R., Dacic, S., 2012. Morphologic and clinicopathologic features of lung squamous cell carcinomas expressing Sox2. *American Journal of Clinical Pathology.* **138,** 712-718.

Bretthauer, M. and Kalager, M., 2013. Principles, effectiveness and caveats in screening for cancer. *The British Journal of Surgery.* **100,** 55-65.

Califano, R., Landi, L., Cappuzzo, F., 2012. Prognostic and predictive value of K-RAS mutations in non-small cell lung cancer. *Drugs.* **72 Suppl 1,** 28-36.

Camps, C., Jantus-Lewintre, E., Cabrera, A., Blasco, A., Sanmartin, E., Gallach, S., Caballero, C., del Pozo, N., Rosell, R., Guijarro, R., Sirera, R., 2011. The identification of KRAS mutations at codon 12 in plasma DNA is not a prognostic factor in advanced non-small cell lung cancer patients. *Lung Cancer (Amsterdam, Netherlands).* **72,** 365-369.

Camps, C., Sirera, R., Bremnes, R.M., Rodenas, V., Blasco, A., Safont, M.J., Garde, J., Juarez, A., Caballero, C., Sanchez, J.J., Taron, M., Rosell, R., 2006. Quantification in the serum of the catalytic fraction of reverse telomerase: a useful prognostic factor in advanced non-small cell lung cancer. *Anticancer Research.* **26,** 4905-4909.

Cancer Genome Atlas Research Network, 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* **511,** 543-550.

Cancer Genome Atlas Research Network, 2012. Comprehensive genomic characterization of squamous cell lung cancers. *Nature.* **489,** 519-525.

Cappuzzo, F., Hirsch, F.R., Rossi, E., Bartolini, S., Ceresoli, G.L., Bemis, L., Haney, J., Witta, S., Danenberg, K., Domenichini, I., Ludovini, V., Magrini, E., Gregorc, V., Doglioni, C., Sidoni, A., Tonato, M., Franklin, W.A., Crino, L., Bunn, P.A.,Jr, Varella-Garcia, M., 2005. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *Journal of the National Cancer Institute.* **97,** 643-655.

Cappuzzo, F., Janne, P.A., Skokan, M., Finocchiaro, G., Rossi, E., Ligorio, C., Zucali, P.A., Terracciano, L., Toschi, L., Roncalli, M., Destro, A., Incarbone, M., Alloisio, M., Santoro, A., Varella-Garcia, M., 2009a. MET increased gene copy number and primary resistance to gefitinib therapy in non-small-cell lung cancer patients. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO.* **20,** 298-304.

Cappuzzo, F., Marchetti, A., Skokan, M., Rossi, E., Gajapathy, S., Felicioni, L., Del Grammastro, M., Sciarrotta, M.G., Buttitta, F., Incarbone, M., Toschi, L., Finocchiaro, G., Destro, A., Terracciano, L., Roncalli, M., Alloisio, M., Santoro, A., Varella-Garcia, M., 2009b. Increased MET gene copy number negatively affects survival of surgically resected non-small-cell lung cancer patients. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **27,** 1667-1674.

Cappuzzo, F., Varella-Garcia, M., Shigematsu, H., Domenichini, I., Bartolini, S., Ceresoli, G.L., Rossi, E., Ludovini, V., Gregorc, V., Toschi, L., Franklin, W.A., Crino, L., Gazdar, A.F., Bunn, P.A.,Jr, Hirsch, F.R., 2005. Increased HER2 gene copy number is associated with response to gefitinib therapy in epidermal growth factor receptor-positive non-small-cell lung cancer patients. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **23,** 5007-5018.

Carriere, C., Seeley, E.S., Goetze, T., Longnecker, D.S., Korc, M., 2007. The Nestin progenitor lineage is the compartment of origin for pancreatic intraepithelial neoplasia. *Proceedings of the National Academy of Sciences of the United States of America.* **104,** 4437-4442.

Catarino, R., Coelho, A., Araujo, A., Gomes, M., Nogueira, A., Lopes, C., Medeiros, R., 2012. Circulating DNA: diagnostic tool and predictive marker for overall survival of NSCLC patients. *PloS One.* **7,** e38559.

Cavanaugh, D., Johnson, E., Price, R.E., Kurie, J., Travis, E.L., Cody, D.D., 2004. In vivo respiratory-gated micro-CT imaging in small-animal oncology models. *Molecular Imaging.* **3,** 55-62.

Chang, A., 2011. Chemotherapy, chemoresistance and the changing treatment landscape for NSCLC. *Lung Cancer (Amsterdam, Netherlands).* **71,** 3-10.

Chang, J.W., Liu, H.P., Hsieh, M.H., Fang, Y.F., Hsieh, M.S., Hsieh, J.J., Chiu, Y.T., Tsai, H.Y., Chen, Y.H., Chen, Y.T., Hsu, H.Y., Chen, Y.T., Tsai, S.F., Chen, Y.R., Hsi, B.L., Huang, S.F., 2008. Increased epidermal growth factor receptor (EGFR) gene copy number is strongly associated with EGFR mutations and adenocarcinoma in non-small cell lung cancers: a chromogenic in situ hybridization study of 182 patients. *Lung Cancer (Amsterdam, Netherlands).* **61,** 328-339.

Chatterjee, P.K., Shakes, L.A., Srivastava, D.K., Garland, D.M., Harewood, K.R., Moore, K.J., Coren, J.S., 2004. Mutually exclusive recombination of wild-type and mutant loxP sites in vivo facilitates transposon-mediated deletions from both ends of genomic DNA in PACs. *Nucleic Acids Research.* **32,** 5668-5676.

Chen, H.J., Mok, T.S., Chen, Z.H., Guo, A.L., Zhang, X.C., Su, J., Wu, Y.L., 2009. Clinicopathologic and molecular features of epidermal growth factor receptor T790M mutation and c-MET amplification in tyrosine kinase inhibitor-resistant Chinese non-small cell lung cancer. *Pathology Oncology Research : POR.* **15,** 651-658.

Chen, P., Li, J., Wang, Y., Zhu, L.R., Hu, Y.M., Tong, X.P., 2013. Detection of FHIT and p16 mRNA deletion in biopsy specimens obtained by bronchoscopy for the diagnosis of lung cancer. *The International Journal of Biological Markers.* 0.

Chen, X.Q., Stroun, M., Magnenat, J.L., Nicod, L.P., Kurt, A.M., Lyautey, J., Lederrey, C., Anker, P., 1996. Microsatellite alterations in plasma DNA of small cell lung cancer patients. *Nature Medicine.* **2,** 1033-1035.

Chen, Z., Zeng, H., Guo, Y., Liu, P., Pan, H., Deng, A., Hu, J., 2010. miRNA-145 inhibits non-small cell lung cancer cell proliferation by targeting c-Myc. *Journal of Experimental and Clinical Cancer Research : CR.* **29,** 151-9966-29-151.

Cheng, C., Omura-Minamisawa, M., Kang, Y., Hara, T., Koike, I., Inoue, T., 2009. Quantification of circulating cell-free DNA in the plasma of cancer patients during radiation therapy. *Cancer Science.* **100,** 303-309.

Cheon, D.J. and Orsulic, S., 2011. Mouse models of cancer. *Annual Review of Pathology.* **6,** 95-119.

Chiappetta, C., Anile, M., Leopizzi, M., Venuta, F., Della Rocca, C., 2013. Use of a new generation of capillary electrophoresis to quantify circulating free DNA in non-small cell lung cancer. *Clinica Chimica Acta; International Journal of Clinical Chemistry.* **425,** 93-96.

Choi, Y.L., Soda, M., Yamashita, Y., Ueno, T., Takashima, J., Nakajima, T., Yatabe, Y., Takeuchi, K., Hamada, T., Haruta, H., Ishikawa, Y., Kimura, H., Mitsudomi, T., Tanio, Y., Mano, H., ALK Lung Cancer Study Group, 2010. EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. *The New England Journal of Medicine.* **363,** 1734-1739.

Chong, I.W., Chang, M.Y., Sheu, C.C., Wang, C.Y., Hwang, J.J., Huang, M.S., Lin, S.R., 2007. Detection of activated K-ras in non-small cell lung cancer by membrane array: a comparison with direct sequencing. *Oncology Reports.* **18,** 17-24.

Churg, A.M., Myers, J.L., Tazelaar, H.D. and Wright, J.L., 2005. *Thurlbeck's Pathology of the Lung.* 3rd ed. New York: Thieme Medical Publishers.

Chyczewski, L., Niklinska, W., Naumnik, W., Chyczewska, E., Niklinski, J., 2001. P16INK4a point mutations and promoter hypermetylation in bronchial preneoplastic lesions. *Folia Histochemica Et Cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society.* **39 Suppl 2,** 33-35.

Ciuleanu, T., Brodowicz, T., Zielinski, C., Kim, J.H., Krzakowski, M., Laack, E., Wu, Y.L., Bover, I., Begbie, S., Tzekova, V., Cucevic, B., Pereira, J.R., Yang, S.H., Madhavan, J., Sugarman, K.P., Peterson, P., John, W.J., Krejcy, K., Belani, C.P., 2009. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomised, double-blind, phase 3 study. *Lancet.* **374,** 1432-1440.

Collado, M., Gil, J., Efeyan, A., Guerra, C., Schuhmacher, A.J., Barradas, M., Benguria, A., Zaballos, A., Flores, J.M., Barbacid, M., Beach, D., Serrano, M., 2005. Tumour biology: senescence in premalignant tumours. *Nature.* **436,** 642.

Cortes-Funes, H., Gomez, C., Rosell, R., Valero, P., Garcia-Giron, C., Velasco, A., Izquierdo, A., Diz, P., Camps, C., Castellanos, D., Alberola, V., Cardenal, F., Gonzalez-Larriba, J.L., Vieitez, J.M., Maeztu, I., Sanchez, J.J., Queralt, C., Mayo, C., Mendez, P., Moran, T., Taron, M., 2005. Epidermal growth factor receptor activating mutations in Spanish gefitinib-treated non-small-cell lung cancer patients. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO.* **16,** 1081-1086.

Czeiger, D., Shaked, G., Eini, H., Vered, I., Belochitski, O., Avriel, A., Ariad, S., Douvdevani, A., 2011. Measurement of circulating cell-free DNA levels by a new simple fluorescent test in patients with primary colorectal cancer. *American Journal of Clinical Pathology.* **135,** 264-270.

Dacic, S., Flanagan, M., Cieply, K., Ramalingam, S., Luketich, J., Belani, C., Yousem, S.A., 2006. Significance of EGFR protein expression and gene amplification in non-small cell lung carcinoma. *American Journal of Clinical Pathology.* **125,** 860-865.

Dankort, D., Filenova, E., Collado, M., Serrano, M., Jones, K., McMahon, M., 2007. A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. *Genes and Development.* **21,** 379-384.

Davies, K.D., Le, A.T., Theodoro, M.F., Skokan, M.C., Aisner, D.L., Berge, E.M., Terracciano, L.M., Cappuzzo, F., Incarbone, M., Roncalli, M., Alloisio, M., Santoro, A., Camidge, D.R., Varella-Garcia, M., Doebele, R.C., 2012. Identifying and targeting ROS1 gene fusions in non-small cell lung cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **18,** 4570-4579.

Dawson, S.J., Tsui, D.W., Murtaza, M., Biggs, H., Rueda, O.M., Chin, S.F., Dunning, M.J., Gale, D., Forshew, T., Mahler-Araujo, B., Rajan, S., Humphray, S., Becq, J., Halsall, D., Wallis, M., Bentley, D., Caldas, C., Rosenfeld, N., 2013. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *The New England Journal of Medicine.* **368,** 1199-1209.

Deeb, K.K., Hohman, C.M., Risch, N.F., Metzger, D.J., Starostik, P., 2015. Routine Clinical Mutation Profiling of Non-Small Cell Lung Cancer Using Next-Generation Sequencing. *Archives of Pathology and Laboratory Medicine.* **139,** 913-921.

Department of Health, 2011. **The likely impact of earlier diagnosis of cancer on costs and benefits to the NHS : summary of an economic modelling project carried out by Frontier Economics on behalf of the Department of Health, the National Awareness and Early Diagnosis Initiative (NAEDI)**. London: .

Devonshire, A.S., Whale, A.S., Gutteridge, A., Jones, G., Cowen, S., Foy, C.A., Huggett, J.F., 2014. Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. *Analytical and Bioanalytical Chemistry.* **406,** 6499-6512.

Dhillon, S.S., Loewen, G., Jayaprakash, V., Reid, M.E., 2013. Lung cancer screening update. *Journal of Carcinogenesis.* **12,** 2-3163.106681. Print 2013.

Diaz, L.A.,Jr and Bardelli, A., 2014. Liquid biopsies: genotyping circulating tumor DNA. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **32,** 579-586.

Diederich, S., Wormanns, D., Semik, M., Thomas, M., Lenzen, H., Roos, N., Heindel, W., 2002. Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. *Radiology.* **222,** 773-781.

Diehl, F., Schmidt, K., Choti, M.A., Romans, K., Goodman, S., Li, M., Thornton, K., Agrawal, N., Sokoll, L., Szabo, S.A., Kinzler, K.W., Vogelstein, B., Diaz, L.A.,Jr, 2008. Circulating mutant DNA to assess tumor dynamics. *Nature Medicine.* **14,** 985-990.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., Fulton, L., Fulton, R.S., Zhang, Q., Wendl, M.C., Lawrence, M.S., Larson, D.E., Chen, K., Dooling, D.J., Sabo, A., Hawes, A.C., Shen, H., Jhangiani, S.N., Lewis, L.R., Hall, O., Zhu, Y., Mathew, T., Ren, Y., Yao, J., Scherer, S.E., Clerc, K., Metcalf, G.A., Ng, B., Milosavljevic, A., Gonzalez-Garay, M.L., Osborne, J.R., Meyer, R., Shi, X., Tang, Y., Koboldt, D.C., Lin, L., Abbott, R., Miner, T.L., Pohl, C., Fewell, G., Haipek, C., Schmidt, H., Dunford-Shore, B.H., Kraja, A., Crosby, S.D., Sawyer, C.S., Vickery, T., Sander, S., Robinson, J., Winckler, W., Baldwin, J., Chirieac, L.R., Dutt, A., Fennell, T., Hanna, M., Johnson, B.E., Onofrio, R.C., Thomas, R.K., Tonon, G., Weir, B.A., Zhao, X., Ziaugra, L., Zody, M.C., Giordano, T., Orringer, M.B., Roth, J.A., Spitz, M.R., Wistuba, I.I., Ozenberger, B., Good, P.J., Chang, A.C., Beer, D.G., Watson, M.A., Ladanyi, M., Broderick, S., Yoshizawa, A., Travis, W.D., Pao, W., Province, M.A., Weinstock, G.M., Varmus, H.E., Gabriel, S.B., Lander, E.S., Gibbs, R.A., Meyerson, M., Wilson, R.K., 2008a. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature.* **455,** 1069-1075.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., Fulton, L., Fulton, R.S., Zhang, Q., Wendl, M.C., Lawrence, M.S., Larson, D.E., Chen, K., Dooling, D.J., Sabo, A., Hawes, A.C., Shen, H., Jhangiani, S.N., Lewis, L.R., Hall, O., Zhu, Y., Mathew, T., Ren, Y., Yao, J., Scherer, S.E., Clerc, K., Metcalf, G.A., Ng, B., Milosavljevic, A., Gonzalez-Garay, M.L., Osborne, J.R., Meyer, R., Shi, X., Tang, Y., Koboldt, D.C., Lin, L., Abbott, R., Miner, T.L., Pohl, C., Fewell, G., Haipek, C., Schmidt, H., Dunford-Shore, B.H., Kraja, A., Crosby, S.D., Sawyer, C.S., Vickery, T., Sander, S., Robinson, J., Winckler, W., Baldwin, J., Chirieac, L.R., Dutt, A., Fennell, T., Hanna, M., Johnson, B.E., Onofrio, R.C., Thomas, R.K., Tonon, G., Weir, B.A., Zhao, X., Ziaugra, L., Zody, M.C., Giordano, T., Orringer, M.B., Roth, J.A., Spitz, M.R., Wistuba, I.I., Ozenberger, B., Good, P.J., Chang, A.C., Beer, D.G., Watson, M.A., Ladanyi, M., Broderick, S., Yoshizawa, A., Travis, W.D., Pao, W., Province, M.A., Weinstock, G.M., Varmus, H.E., Gabriel, S.B., Lander, E.S., Gibbs, R.A., Meyerson, M., Wilson, R.K., 2008b. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature.* **455,** 1069-1075.

Doria-Rose, V.P., Marcus, P.M., Szabo, E., Tockman, M.S., Melamed, M.R., Prorok, P.C., 2009. Randomized controlled trials of the efficacy of lung cancer

screening by sputum cytology revisited: a combined mortality analysis from the Johns Hopkins Lung Project and the Memorial Sloan-Kettering Lung Study. *Cancer.* **115,** 5007-5017.

DuPage, M., Dooley, A.L., Jacks, T., 2009. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols.* **4,** 1064-1072.

Dutt, A., Ramos, A.H., Hammerman, P.S., Mermel, C., Cho, J., Sharifnia, T., Chande, A., Tanaka, K.E., Stransky, N., Greulich, H., Gray, N.S., Meyerson, M., 2011. Inhibitor-sensitive FGFR1 amplification in human non-small cell lung cancer. *PloS One.* **6,** e20351.

eBioscience, A., 2016*. QuantiGene® DNA Copy Number Variation Assay-Data and Specifications.* [online]. Available at: http://www.ebioscience.com/application/gene-expression/DNA-Copy-Data-Specifications.htm [accessed 09/17 2016].

Edge, S.B., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L. and Trotti, A., 2010. Part IV *AJCC Cancer Staging Manual.* 7th ed. New York: Springer. 253-270.

Elshimali, Y.I., Khaddour, H., Sarkissyan, M., Wu, Y., Vadgama, J.V., 2013. The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients. *International Journal of Molecular Sciences.* **14,** 18925-18958.

Esposito, D., Gillette, W.K., Hartley, J.L., 2003. Blocking oligonucleotides improve sequencing through inverted repeats. *BioTechniques.* **35,** 914-6, 918, 920.

Faca, V.M., Song, K.S., Wang, H., Zhang, Q., Krasnoselsky, A.L., Newcomb, L.F., Plentz, R.R., Gurumurthy, S., Redston, M.S., Pitteri, S.J., Pereira-Faca, S.R., Ireton, R.C., Katayama, H., Glukhova, V., Phanstiel, D., Brenner, D.E., Anderson, M.A., Misek, D., Scholler, N., Urban, N.D., Barnett, M.J., Edelstein, C., Goodman, G.E., Thornquist, M.D., McIntosh, M.W., DePinho, R.A., Bardeesy, N., Hanash, S.M., 2008. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Medicine.* **5,** e123.

Fishback, N.F., Travis, W.D., Moran, C.A., Guinee, D.G.,Jr, McCarthy, W.F., Koss, M.N., 1994. Pleomorphic (spindle/giant cell) carcinoma of the lung. A clinicopathologic correlation of 78 cases. *Cancer.* **73,** 2936-2945.

Fisher, G.H., Wellen, S.L., Klimstra, D., Lenczowski, J.M., Tichelaar, J.W., Lizak, M.J., Whitsett, J.A., Koretsky, A., Varmus, H.E., 2001. Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras

transgene in the presence and absence of tumor suppressor genes. *Genes and Development.* **15,** 3249-3262.

Flehinger, B.J., Melamed, M.R., Zaman, M.B., Heelan, R.T., Perchick, W.B., Martini, N., 1984. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Memorial Sloan-Kettering study. *The American Review of Respiratory Disease.* **130,** 555-560.

Fleischhacker, M., Schmidt, B. 2008. Cell-free DNA resuscitated for tumor testing. *Nature Medicine.* **14**, 914-915

Folkert, M.R. and Timmerman, R., 2015. Review of treatment options for oligometastatic non-small cell lung cancer. *Clinical Advances in Hematology and Oncology : HandO.* **13,** 186-193.

Fontana, R.S., Sanderson, D.R., Woolner, L.B., Taylor, W.F., Miller, W.E., Muhm, J.R., Bernatz, P.E., Payne, W.S., Pairolero, P.C., Bergstralh, E.J., 1991. Screening for lung cancer. A critique of the Mayo Lung Project. *Cancer.* **67,** 1155-1164.

Forde, P.M. and Rudin, C.M., 2012. Crizotinib in the treatment of non-small-cell lung cancer. *Expert Opinion on Pharmacotherapy.* **13,** 1195-1201.

Forshew, T., Murtaza, M., Parkinson, C., Gale, D., Tsui, D.W., Kaper, F., Dawson, S.J., Piskorz, A.M., Jimenez-Linan, M., Bentley, D., Hadfield, J., May, A.P., Caldas, C., Brenton, J.D., Rosenfeld, N., 2012. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Science Translational Medicine.* **4,** 136ra68.

Franklin, W.A., 2000. Pathology of lung cancer. *Journal of Thoracic Imaging.* **15,** 3-12.

Frese, K.K. and Tuveson, D.A., 2007. Maximizing mouse cancer models. *Nature Reviews.Cancer.* **7,** 645-658.

Frost, J.K., Ball, W.C.,Jr, Levin, M.L., Tockman, M.S., Baker, R.R., Carter, D., Eggleston, J.C., Erozan, Y.S., Gupta, P.K., Khouri, N.F., 1984. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. *The American Review of Respiratory Disease.* **130,** 549-554.

Garcia-Olmo, D., Garcia-Olmo, D.C., Dominguez-Berzosa, C., Guadalajara, H., Vega, L., Garcia-Arranz, M., 2012. Oncogenic transformation induced by cell-free nucleic acids circulating in plasma (genometastasis) remains after the surgical resection of the primary tumor: a pilot study. *Expert Opinion on Biological Therapy.* **12 Suppl 1,** S61-8.

Garcia-Olmo, D., Garcia-Olmo, D.C., Ontanon, J., Martinez, E., 2000. Horizontal transfer of DNA and the "genometastasis hypothesis". *Blood.* **95,** 724-725.

Garcia-Olmo, D., Garcia-Olmo, D.C., Ontanon, J., Martinez, E., Vallejo, M., 1999a. Tumor DNA circulating in the plasma might play a role in metastasis. The hypothesis of the genometastasis. *Histology and Histopathology.* **14,** 1159-1164.

Garcia-Olmo, D., Ontanon, J., Garcia-Olmo, D.C., Atienzar, M., Vallejo, M., 1999b. Detection of genomically-tagged cancer cells in different tissues at different stages of tumor development: lack of correlation with the formation of metastasis. *Cancer Letters.* **140,** 11-20.

Garcia-Olmo, D.C., Dominguez, C., Garcia-Arranz, M., Anker, P., Stroun, M., Garcia-Verdugo, J.M., Garcia-Olmo, D., 2010. Cell-free nucleic acids circulating in the plasma of colorectal cancer patients induce the oncogenic transformation of susceptible cultured cells. *Cancer Research.* **70,** 560-567.

Garcia-Olmo, D.C., Ruiz-Piqueras, R., Garcia-Olmo, D., 2004. Circulating nucleic acids in plasma and serum (CNAPS) and its relation to stem cells and cancer metastasis: state of the issue. *Histology and Histopathology.* **19,** 575-583.

Gautschi, O., Bigosch, C., Huegli, B., Jermann, M., Marx, A., Chasse, E., Ratschiller, D., Weder, W., Joerger, M., Betticher, D.C., Stahel, R.A., Ziegler, A., 2004. Circulating deoxyribonucleic Acid as prognostic marker in non-small-cell lung cancer patients undergoing chemotherapy. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **22,** 4157-4164.

Gautschi, O., Huegli, B., Ziegler, A., Gugger, M., Heighway, J., Ratschiller, D., Mack, P.C., Gumerlock, P.H., Kung, H.J., Stahel, R.A., Gandara, D.R., Betticher, D.C., 2007. Origin and prognostic value of circulating KRAS mutations in lung cancer patients. *Cancer Letters.* **254,** 265-273.

George, J., Lim, J.S., Jang, S.J., Cun, Y., Ozretic, L., Kong, G., Leenders, F., Lu, X., Fernandez-Cuesta, L., Bosco, G., Muller, C., Dahmen, I., Jahchan, N.S., Park, K.S., Yang, D., Karnezis, A.N., Vaka, D., Torres, A., Wang, M.S., Korbel, J.O., Menon, R., Chun, S.M., Kim, D., Wilkerson, M., Hayes, N., Engelmann, D., Putzer, B., Bos, M., Michels, S., Vlasic, I., Seidel, D., Pinther, B., Schaub, P., Becker, C., Altmuller, J., Yokota, J., Kohno, T., Iwakawa, R., Tsuta, K., Noguchi, M., Muley, T., Hoffmann, H., Schnabel, P.A., Petersen, I., Chen, Y., Soltermann, A., Tischler, V., Choi, C.M., Kim, Y.H., Massion, P.P., Zou, Y., Jovanovic, D., Kontic, M., Wright, G.M., Russell, P.A., Solomon, B., Koch, I., Lindner, M., Muscarella, L.A., la Torre, A., Field, J.K., Jakopovic, M., Knezevic, J., Castanos-Velez, E., Roz, L., Pastorino, U., Brustugun, O.T., Lund-Iversen,

M., Thunnissen, E., Kohler, J., Schuler, M., Botling, J., Sandelin, M., Sanchez-Cespedes, M., Salvesen, H.B., Achter, V., Lang, U., Bogus, M., Schneider, P.M., Zander, T., Ansen, S., Hallek, M., Wolf, J., Vingron, M., Yatabe, Y., Travis, W.D., Nurnberg, P., Reinhardt, C., Perner, S., Heukamp, L., Buttner, R., Haas, S.A., Brambilla, E., Peifer, M., Sage, J., Thomas, R.K., 2015. Comprehensive genomic profiles of small cell lung cancer. *Nature.* **524,** 47-53.

Gevensleben, H., Garcia-Murillas, I., Graeser, M.K., Schiavon, G., Osin, P., Parton, M., Smith, I.E., Ashworth, A., Turner, N.C., 2013. Noninvasive detection of HER2 amplification with plasma DNA digital PCR. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **19,** 3276-3284.

Giacona, M.B., Ruben, G.C., Iczkowski, K.A., Roos, T.B., Porter, D.M., Sorenson, G.D., 1998. Cell-free DNA in human blood plasma: length measurements in patients with pancreatic cancer and healthy controls. *Pancreas.* **17,** 89-97.

Gidekel Friedlander, S.Y., Chu, G.C., Snyder, E.L., Girnius, N., Dibelius, G., Crowley, D., Vasile, E., DePinho, R.A., Jacks, T., 2009. Context-dependent transformation of adult pancreatic cells by oncogenic K-Ras. *Cancer Cell.* **16,** 379-389.

Go, H., Jeon, Y.K., Park, H.J., Sung, S.W., Seo, J.W., Chung, D.H., 2010. High MET gene copy number leads to shorter survival in patients with non-small cell lung cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **5,** 305-313.

Gold, B., Cankovic, M., Furtado, L.V., Meier, F., Gocke, C.D., 2015. Do Circulating Tumor Cells, Exosomes, and Circulating Tumor Nucleic Acids Have Clinical Utility?: A Report of the Association for Molecular Pathology. *The Journal of Molecular Diagnostics : JMD.*

Gonzalez, R., Silva, J.M., Sanchez, A., Dominguez, G., Garcia, J.M., Chen, X.Q., Stroun, M., Provencio, M., Espana, P., Anker, P., Bonilla, F., 2000. Microsatellite alterations and TP53 mutations in plasma DNA of small-cell lung cancer patients: follow-up study and prognostic significance. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO.* **11,** 1097-1104.

Gopal, M., Abdullah, S.E., Grady, J.J., Goodwin, J.S., 2010. Screening for lung cancer with low-dose computed tomography: a systematic review and meta-analysis of the baseline findings of randomized controlled trials. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **5,** 1233-1239.

Gorges, T.M., Schiller, J., Schmitz, A., Schuetzmann, D., Schatz, C., Zollner, T.M., Krahn, T., von Ahsen, O., 2012. Cancer therapy monitoring in xenografts by quantitative analysis of circulating tumor DNA. *Biomarkers : Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals.* **17,** 498-506.

Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N.D., Kanchi, K.L., Maher, C.A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Bose, R., Ornitz, D., Xiong, D., You, M., Dooling, D.J., Watson, M., Mardis, E.R., Wilson, R.K., 2012a. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell.* **150,** 1121-1134.

Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N.D., Kanchi, K.L., Maher, C.A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Bose, R., Ornitz, D., Xiong, D., You, M., Dooling, D.J., Watson, M., Mardis, E.R., Wilson, R.K., 2012b. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell.* **150,** 1121-1134.

Guescini, M., Sisti, D., Rocchi, M.B., Stocchi, L., Stocchi, V., 2008. A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition. *BMC Bioinformatics.* **9,** 326-2105-9-326.

Guttery, D.S., Page, K., Hills, A., Woodley, L., Marchese, S.D., Rghebi, B., Hastings, R.K., Luo, J., Pringle, J.H., Stebbing, J., Coombes, R.C., Ali, S., Shaw, J.A., 2015. Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. *Clinical Chemistry.* **61,** 974-982.

Hagemann, I.S., Devarakonda, S., Lockwood, C.M., Spencer, D.H., Guebert, K., Bredemeyer, A.J., Al-Kateb, H., Nguyen, T.T., Duncavage, E.J., Cottrell, C.E., Kulkarni, S., Nagarajan, R., Seibert, K., Baggstrom, M., Waqar, S.N., Pfeifer, J.D., Morgensztern, D., Govindan, R., 2015. Clinical next-generation sequencing in patients with non-small cell lung cancer. *Cancer.* **121,** 631-639.

Haines, B.B., Bettano, K.A., Chenard, M., Sevilla, R.S., Ware, C., Angagaw, M.H., Winkelmann, C.T., Tong, C., Reilly, J.F., Sur, C., Zhang, W., 2009. A quantitative volumetric micro-computed tomography method to analyze lung tumors in genetically engineered mouse models. *Neoplasia (New York, N.Y.).* **11,** 39-47.

Hamakawa, T., Kukita, Y., Kurokawa, Y., Miyazaki, Y., Takahashi, T., Yamasaki, M., Miyata, H., Nakajima, K., Taniguchi, K., Takiguchi, S., Mori, M., Doki, Y., Kato, K., 2015. Monitoring gastric cancer progression with circulating tumour DNA. *British Journal of Cancer.* **112,** 352-356.

Hashad, D., Sorour, A., Ghazal, A., Talaat, I., 2012. Free circulating tumor DNA as a diagnostic marker for breast cancer. *Journal of Clinical Laboratory Analysis.* **26,** 467-472.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., Ira, G., 2009. Mechanisms of change in gene copy number. *Nature Reviews.Genetics.* **10,** 551-564.

Heinmoller, P., Gross, C., Beyser, K., Schmidtgen, C., Maass, G., Pedrocchi, M., Ruschoff, J., 2003. HER2 status in non-small cell lung cancer: results from patient screening for enrollment to a phase II study of herceptin. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **9,** 5238-5243.

Heist, R.S., Mino-Kenudson, M., Sequist, L.V., Tammireddy, S., Morrissey, L., Christiani, D.C., Engelman, J.A., Iafrate, A.J., 2012. FGFR1 amplification in squamous cell carcinoma of the lung. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **7,** 1775-1780.

Heitzer, E., Ulz, P., Belic, J., Gutschi, S., Quehenberger, F., Fischereder, K., Benezeder, T., Auer, M., Pischler, C., Mannweiler, S., Pichler, M., Eisner, F., Haeusler, M., Riethdorf, S., Pantel, K., Samonigg, H., Hoefler, G., Augustin, H., Geigl, J.B., Speicher, M.R., 2013. Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. *Genome Medicine.* **5,** 30.

Henschke, C.I., McCauley, D.I., Yankelevitz, D.F., Naidich, D.P., McGuinness, G., Miettinen, O.S., Libby, D., Pasmantier, M., Koizumi, J., Altorki, N., Smith, J.P., 2001. Early lung cancer action project: a summary of the findings on baseline screening. *The Oncologist.* **6,** 147-152.

Henschke, C.I., McCauley, D.I., Yankelevitz, D.F., Naidich, D.P., McGuinness, G., Miettinen, O.S., Libby, D.M., Pasmantier, M.W., Koizumi, J., Altorki, N.K., Smith, J.P., 1999. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet.* **354,** 99-105.

Herrera, L.J., Raja, S., Gooding, W.E., El-Hefnawy, T., Kelly, L., Luketich, J.D., Godfrey, T.E., 2005. Quantitative analysis of circulating plasma DNA as a tumor marker in thoracic malignancies. *Clinical Chemistry.* **51,** 113-118.

Hertoghs, K.M., Ellis, J.H., Catchpole, I.R., 2003. Use of locked nucleic acid oligonucleotides to add functionality to plasmid DNA. *Nucleic Acids Research.* **31,** 5817-5830.

Hingorani, S.R., Petricoin, E.F., Maitra, A., Rajapakse, V., King, C., Jacobetz, M.A., Ross, S., Conrads, T.P., Veenstra, T.D., Hitt, B.A., Kawaguchi, Y., Johann, D., Liotta, L.A., Crawford, H.C., Putt, M.E., Jacks, T., Wright, C.V.,

Hruban, R.H., Lowy, A.M., Tuveson, D.A., 2003. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell.* **4,** 437-450.

Hirsch, F.R., Varella-Garcia, M., Bunn, P.A.,Jr, Di Maria, M.V., Veve, R., Bremmes, R.M., Baron, A.E., Zeng, C., Franklin, W.A., 2003. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **21,** 3798-3807.

Hirsch, F.R., Varella-Garcia, M., Bunn, P.A.,Jr, Franklin, W.A., Dziadziuszko, R., Thatcher, N., Chang, A., Parikh, P., Pereira, J.R., Ciuleanu, T., von Pawel, J., Watkins, C., Flannery, A., Ellison, G., Donald, E., Knight, L., Parums, D., Botwood, N., Holloway, B., 2006. Molecular predictors of outcome with gefitinib in a phase III placebo-controlled study in advanced non-small-cell lung cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **24,** 5034-5042.

Hirsch, F.R., Varella-Garcia, M., Franklin, W.A., Veve, R., Chen, L., Helfrich, B., Zeng, C., Baron, A., Bunn, P.A.,Jr, 2002. Evaluation of HER-2/neu gene amplification and protein expression in non-small cell lung carcinomas. *British Journal of Cancer.* **86,** 1449-1456.

Hirsch, F.R., Varella-Garcia, M., McCoy, J., West, H., Xavier, A.C., Gumerlock, P., Bunn, P.A.,Jr, Franklin, W.A., Crowley, J., Gandara, D.R., Southwest Oncology Group, 2005. Increased epidermal growth factor receptor gene copy number detected by fluorescence in situ hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **23,** 6838-6845.

Huang, F.T., Yu, K., Balter, B.B., Selsing, E., Oruc, Z., Khamlichi, A.A., Hsieh, C.L., Lieber, M.R., 2007. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region. *Molecular and Cellular Biology.* **27,** 5921-5932.

Hung, K.E., Faca, V., Song, K., Sarracino, D.A., Richard, L.G., Krastins, B., Forrester, S., Porter, A., Kunin, A., Mahmood, U., Haab, B.B., Hanash, S.M., Kucherlapati, R., 2009. Comprehensive proteome analysis of an Apc mouse model uncovers proteins associated with intestinal tumorigenesis. *Cancer Prevention Research (Philadelphia, Pa.).* **2,** 224-233.

Hung, K.E., Kho, A.T., Sarracino, D., Richard, L.G., Krastins, B., Forrester, S., Haab, B.B., Kohane, I.S., Kucherlapati, R., 2006. Mass spectrometry-based study of the plasma proteome in a mouse intestinal tumor model. *Journal of Proteome Research.* **5,** 1866-1878.

Hurwitz, J.L., McCoy, F., Scullin, P., Fennell, D.A., 2009. New advances in the second-line treatment of small cell lung cancer. *The Oncologist.* **14,** 986-994.

Hussenet, T., Dali, S., Exinger, J., Monga, B., Jost, B., Dembele, D., Martinet, N., Thibault, C., Huelsken, J., Brambilla, E., du Manoir, S., 2010a. SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. *PloS One.* **5,** e8960.

Hussenet, T., Dali, S., Exinger, J., Monga, B., Jost, B., Dembele, D., Martinet, N., Thibault, C., Huelsken, J., Brambilla, E., du Manoir, S., 2010b. SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. *PloS One.* **5,** e8960.

Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., Sougnez, C., Auclair, D., Lawrence, M.S., Stojanov, P., Cibulskis, K., Choi, K., de Waal, L., Sharifnia, T., Brooks, A., Greulich, H., Banerji, S., Zander, T., Seidel, D., Leenders, F., Ansen, S., Ludwig, C., Engel-Riedel, W., Stoelben, E., Wolf, J., Goparju, C., Thompson, K., Winckler, W., Kwiatkowski, D., Johnson, B.E., Janne, P.A., Miller, V.A., Pao, W., Travis, W.D., Pass, H.I., Gabriel, S.B., Lander, E.S., Thomas, R.K., Garraway, L.A., Getz, G., Meyerson, M., 2012. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell.* **150,** 1107-1120.

Iranmanesh, S.M. and Guo, N.L., 2014. Integrated DNA Copy Number and Gene Expression Regulatory Network Analysis of Non-small Cell Lung Cancer Metastasis. *Cancer Informatics.* **13,** 13-23.

Isozaki, H., Takigawa, N., Kiura, K., 2015. Mechanisms of Acquired Resistance to ALK Inhibitors and the Rationale for Treating ALK-positive Lung Cancer. *Cancers.* **7,** 763-783.

Jackman, D.M. and Johnson, B.E., 2005. Small-cell lung cancer. *Lancet.* **366,** 1385-1396.

Jackson, E.L., Olive, K.P., Tuveson, D.A., Bronson, R., Crowley, D., Brown, M., Jacks, T., 2005. The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Research.* **65,** 10280-10288.

Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., Tuveson, D.A., 2001. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes and Development.* **15,** 3243-3248.

Jahr, S., Hentze, H., Englisch, S., Hardt, D., Fackelmayer, F.O., Hesch, R.D., Knippers, R., 2001. DNA fragments in the blood plasma of cancer patients:

quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Research.* **61,** 1659-1665.

Jeanmart, M., Lantuejoul, S., Fievet, F., Moro, D., Sturm, N., Brambilla, C., Brambilla, E., 2003. Value of immunohistochemical markers in preinvasive bronchial lesions in risk assessment of lung cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **9,** 2195-2203.

Jeremy George, P., Banerjee, A.K., Read, C.A., O'Sullivan, C., Falzon, M., Pezzella, F., Nicholson, A.G., Shaw, P., Laurent, G., Rabbitts, P.H., 2007. Surveillance for the detection of early lung cancer in patients with bronchial dysplasia. *Thorax.* **62,** 43-50.

Ji, H., Ramsey, M.R., Hayes, D.N., Fan, C., McNamara, K., Kozlowski, P., Torrice, C., Wu, M.C., Shimamura, T., Perera, S.A., Liang, M.C., Cai, D., Naumov, G.N., Bao, L., Contreras, C.M., Li, D., Chen, L., Krishnamurthy, J., Koivunen, J., Chirieac, L.R., Padera, R.F., Bronson, R.T., Lindeman, N.I., Christiani, D.C., Lin, X., Shapiro, G.I., Janne, P.A., Johnson, B.E., Meyerson, M., Kwiatkowski, D.J., Castrillon, D.H., Bardeesy, N., Sharpless, N.E., Wong, K.K., 2007. LKB1 modulates lung cancer differentiation and metastasis. *Nature.* **448,** 807-810.

Ji, M., Guan, H., Gao, C., Shi, B., Hou, P., 2011. Highly frequent promoter methylation and PIK3CA amplification in non-small cell lung cancer (NSCLC). *BMC Cancer.* **11,** 147-2407-11-147.

Jiang, P., Chan, C.W., Chan, K.C., Cheng, S.H., Wong, J., Wong, V.W., Wong, G.L., Chan, S.L., Mok, T.S., Chan, H.L., Lai, P.B., Chiu, R.W., Lo, Y.M., 2015. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proceedings of the National Academy of Sciences of the United States of America.* **112,** E1317-25.

Jiang, T., Zhai, C., Su, C., Ren, S., Zhou, C., 2016. The diagnostic value of circulating cell free DNA quantification in non-small cell lung cancer: A systematic review with meta-analysis. *Lung Cancer (Amsterdam, Netherlands).* **100,** 63-70.

Jing, C.W., Wang, Z., Cao, H.X., Ma, R., Wu, J.Z., 2014. High resolution melting analysis for epidermal growth factor receptor mutations in formalin-fixed paraffin-embedded tissue and plasma free DNA from non-small cell lung cancer patients. *Asian Pacific Journal of Cancer Prevention : APJCP.* **14,** 6619-6623.

Johnson, B.E., Brennan, J.F., Ihde, D.C., Gazdar, A.F., 1992. myc family DNA amplification in tumors and tumor cell lines from patients with small-cell lung cancer. *Journal of the National Cancer Institute.Monographs.* **(13),** 39-43.

Johnson, J.L., Pillai, S., Chellappan, S.P., 2012. Genetic and biochemical alterations in non-small cell lung cancer. *Biochemistry Research International.* **2012,** 940405.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R.T., Crowley, D., Tuveson, D.A., Jacks, T., 2001. Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature.* **410,** 1111-1116.

Jonsson, S., Varella-Garcia, M., Miller, Y.E., Wolf, H.J., Byers, T., Braudrick, S., Kiatsimkul, P., Lewis, M., Kennedy, T.C., Keith, R.L., Bjornsson, J., McWilliams, A., Lam, S., Hirsch, F.R., Franklin, W.A., 2008. Chromosomal aneusomy in bronchial high-grade lesions is associated with invasive lung cancer. *American Journal of Respiratory and Critical Care Medicine.* **177,** 342-347.

Jorge, S.E., Kobayashi, S.S., Costa, D.B., 2014. Epidermal growth factor receptor (EGFR) mutations in lung cancer: preclinical and clinical data. *Brazilian Journal of Medical and Biological Research = Revista Brasileira De Pesquisas Medicas e Biologicas / Sociedade Brasileira De Biofisica ...[Et Al.].* **0,** 0.

Ju, H.L., Calvisi, D.F., Moon, H., Baek, S., Ribback, S., Dombrowski, F., Cho, K.J., Chung, S.I., Han, K.H., Ro, S.W., 2015. Transgenic mouse model expressing P53(R172H), luciferase, EGFP, and KRAS(G12D) in a single open reading frame for live imaging of tumor. *Scientific Reports.* **5,** 8053.

Junker, K., Stachetzki, U., Rademacher, D., Linder, A., Macha, H.N., Heinecke, A., Muller, K.M., Thomas, M., 2005. HER2/neu expression and amplification in non-small cell lung cancer prior to and after neoadjuvant therapy. *Lung Cancer (Amsterdam, Netherlands).* **48,** 59-67.

Kallioniemi, O.P., Kallioniemi, A., Kurisu, W., Thor, A., Chen, L.C., Smith, H.S., Waldman, F.M., Pinkel, D., Gray, J.W., 1992. ERBB2 amplification in breast cancer analyzed by fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences of the United States of America.* **89,** 5321-5325.

Kamat, A.A., Bischoff, F.Z., Dang, D., Baldwin, M.F., Han, L.Y., Lin, Y.G., Merritt, W.M., Landen, C.N.,Jr, Lu, C., Gershenson, D.M., Simpson, J.L., Sood, A.K., 2006. Circulating cell-free DNA: a novel biomarker for response to therapy in ovarian carcinoma. *Cancer Biology and Therapy.* **5,** 1369-1374.

Kaneko, M., Eguchi, K., Ohmatsu, H., Kakinuma, R., Naruke, T., Suemasu, K., Moriyama, N., 1996. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. *Radiology.* **201,** 798-802.

Kawano, O., Sasaki, H., Endo, K., Suzuki, E., Haneda, H., Yukiue, H., Kobayashi, Y., Yano, M., Fujii, Y., 2006. PIK3CA mutation status in Japanese lung cancer patients. *Lung Cancer (Amsterdam, Netherlands).* **54,** 209-215.

Kawano, O., Sasaki, H., Okuda, K., Yukiue, H., Yokoyama, T., Yano, M., Fujii, Y., 2007. PIK3CA gene amplification in Japanese non-small cell lung cancer. *Lung Cancer (Amsterdam, Netherlands).* **58,** 159-160.

Kelly-Spratt, K.S., Kasarda, A.E., Igra, M., Kemp, C.J., 2008. A mouse model repository for cancer biomarker discovery. *Journal of Proteome Research.* **7,** 3613-3618.

Kerr, K.M., 2001. Pulmonary preinvasive neoplasia. *Journal of Clinical Pathology.* **54,** 257-271.

Khan, S., Coulson, J.M., Woll, P.J., 2004. Genetic abnormalities in plasma DNA of patients with lung cancer and other respiratory diseases. *International Journal of Cancer.Journal International Du Cancer.* **110,** 891-895.

Kim, C.F., Jackson, E.L., Woolfenden, A.E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R.T., Jacks, T., 2005. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell.* **121,** 823-835.

Kim, H.R., Kim, D.J., Kang, D.R., Lee, J.G., Lim, S.M., Lee, C.Y., Rha, S.Y., Bae, M.K., Lee, Y.J., Kim, S.H., Ha, S.J., Soo, R.A., Chung, K.Y., Kim, J.H., Lee, J.H., Shim, H.S., Cho, B.C., 2013a. Fibroblast growth factor receptor 1 gene amplification is associated with poor survival and cigarette smoking dosage in patients with resected squamous cell lung cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **31,** 731-737.

Kim, H.R., Lee, S.Y., Hyun, D.S., Lee, M.K., Lee, H.K., Choi, C.M., Yang, S.H., Kim, Y.C., Lee, Y.C., Kim, S.Y., Jang, S.H., Lee, J.C., Lee, K.Y., 2013b. Detection of EGFR mutations in circulating free DNA by PNA-mediated PCR clamping. *Journal of Experimental and Clinical Cancer Research : CR.* **32,** 50-9966-32-50.

Kim, K., Shin, D.G., Park, M.K., Baik, S.H., Kim, T.H., Kim, S., Lee, S., 2014. Circulating cell-free DNA as a promising biomarker in patients with gastric cancer: diagnostic validity and significant reduction of cfDNA after surgical resection. *Annals of Surgical Treatment and Research.* **86,** 136-142.

Kinugasa, H., Nouso, K., Tanaka, T., Miyahara, K., Morimoto, Y., Dohi, C., Matsubara, T., Okada, H., Yamamoto, K., 2015. Droplet digital PCR measurement of HER2 in patients with gastric cancer. *British Journal of Cancer.* **112,** 1652-1655.

Kirkizlar, E., Zimmermann, B., Constantin, T., Swenerton, R., Hoang, B., Wayham, N., Babiarz, J.E., Demko, Z., Pelham, R.J., Kareht, S., Simon, A.L., Jinnett, K.N., Rabinowitz, M., Sigurjonsson, S., Hill, M., 2015. Detection of Clonal and Subclonal Copy-Number Variants in Cell-Free DNA from Patients

with Breast Cancer Using a Massively Multiplexed PCR Methodology. *Translational Oncology.* **8,** 407-416.

Kirsch, D.G., Grimm, J., Guimaraes, A.R., Wojtkiewicz, G.R., Perez, B.A., Santiago, P.M., Anthony, N.K., Forbes, T., Doppke, K., Weissleder, R., Jacks, T., 2010. Imaging primary lung cancers in mice to study radiation biology. *International Journal of Radiation Oncology, Biology, Physics.* **76,** 973-977.

Kitamura, H., Yazawa, T., Okudela, K., Shimoyamada, H., Sato, H., 2008. Molecular and Genetic Pathogenesis of Lung Cancer: Differences Between Small-Cell and Non-Small-Cell Carcinomas. *The Open Pathology Journa.* **2,** 106-114.

Kitamura, H., Kameda, Y., Ito, T., Hayashi, H., 1999. Atypical adenomatous hyperplasia of the lung. Implications for the pathogenesis of peripheral lung adenocarcinoma. *American Journal of Clinical Pathology.* **111,** 610-622.

Kobayashi, S., Boggon, T.J., Dayaram, T., Janne, P.A., Kocher, O., Meyerson, M., Johnson, B.E., Eck, M.J., Tenen, D.G., Halmos, B., 2005. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *The New England Journal of Medicine.* **352,** 786-792.

Kohno, T., Ichikawa, H., Totoki, Y., Yasuda, K., Hiramoto, M., Nammo, T., Sakamoto, H., Tsuta, K., Furuta, K., Shimada, Y., Iwakawa, R., Ogiwara, H., Oike, T., Enari, M., Schetter, A.J., Okayama, H., Haugen, A., Skaug, V., Chiku, S., Yamanaka, I., Arai, Y., Watanabe, S., Sekine, I., Ogawa, S., Harris, C.C., Tsuda, H., Yoshida, T., Yokota, J., Shibata, T., 2012. KIF5B-RET fusions in lung adenocarcinoma. *Nature Medicine.* **18,** 375-377.

Kontanis, E.J. and Reed, F.A., 2006. Evaluation of real-time PCR amplification efficiencies to detect PCR inhibitors. *Journal of Forensic Sciences.* **51,** 795-804.

Korenkova, V., Scott, J., Novosadova, V., Jindrichova, M., Langerova, L., Svec, D., Sidova, M., Sjoback, R., 2015. Pre-amplification in the context of high-throughput qPCR gene expression experiment. *BMC Molecular Biology.* **16,** 5-015-0033-9.

Kucherlapati, R., 2012. Genetically modified mouse models for biomarker discovery and preclinical drug testing. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **18,** 625-630.

Kulka, J., Tokes, A.M., Kaposi-Novak, P., Udvarhelyi, N., Keller, A., Schaff, Z., 2006. Detection of HER-2/neu gene amplification in breast carcinomas using quantitative real-time PCR - a comparison with immunohistochemical and FISH results. *Pathology Oncology Research : POR.* **12,** 197-204.

Kumar, S., Guleria, R., Singh, V., Bharti, A.C., Mohan, A., Das, B.C., 2010. Plasma DNA level in predicting therapeutic efficacy in advanced nonsmall cell lung cancer. *The European Respiratory Journal.* **36,** 885-892.

Ladanyi, M. and Pao, W., 2008. Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.* **21 Suppl 2,** S16-22.

Lalwani, K., Giddabasappa, A., Li, D., Olson, P., Simmons, B., Shojaei, F., Van Arsdale, T., Christensen, J., Jackson-Fisher, A., Wong, A., Lappin, P.B., Eswaraka, J., 2013. Contrast agents for quantitative microCT of lung tumors in mice. *Comparative Medicine.* **63,** 482-490.

Lamy, P.J., Fina, F., Bascoul-Mollevi, C., Laberenne, A.C., Martin, P.M., Ouafik, L., Jacot, W., 2011. Quantification and clinical relevance of gene amplification at chromosome 17q12-q21 in human epidermal growth factor receptor 2-amplified breast cancers. *Breast Cancer Research : BCR.* **13,** R15.

Lantuejoul, S., Salameire, D., Salon, C., Brambilla, E., 2009. Pulmonary preneoplasia--sequential molecular carcinogenetic events. *Histopathology.* **54,** 43-54.

Laurie, S.A. and Goss, G.D., 2013. Role of epidermal growth factor receptor inhibitors in epidermal growth factor receptor wild-type non-small-cell lung cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **31,** 1061-1069.

Lee, W., Jiang, Z., Liu, J., Haverty, P.M., Guan, Y., Stinson, J., Yue, P., Zhang, Y., Pant, K.P., Bhatt, D., Ha, C., Johnson, S., Kennemer, M.I., Mohan, S., Nazarenko, I., Watanabe, C., Sparks, A.B., Shames, D.S., Gentleman, R., de Sauvage, F.J., Stern, H., Pandita, A., Ballinger, D.G., Drmanac, R., Modrusan, Z., Seshagiri, S., Zhang, Z., 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature.* **465,** 473-477.

Leon, S.A., Shapiro, B., Sklaroff, D.M., Yaros, M.J., 1977. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Research.* **37,** 646-650.

Lester, J.F., Macbeth, F.R., Toy, E., Coles, B., 2006. Palliative radiotherapy regimens for non-small cell lung cancer. *The Cochrane Database of Systematic Reviews.* **(4),** CD002143.

Li, B.T., Drilon, A., Johnson, M.L., Hsu, M., Sima, C.S., McGinn, C., Sugita, H., Kris, M.G., Azzoli, C.G., 2016. A prospective study of total plasma cell-free DNA as a predictive biomarker for response to systemic therapy in patients with advanced non-small-cell lung cancers. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO.* **27,** 154-159.

Liang, Z., Zhang, J., Zeng, X., Gao, J., Wu, S., Liu, T., 2010. Relationship between EGFR expression, copy number and mutation in lung adenocarcinomas. *BMC Cancer.* **10,** 376-2407-10-376.

Lipson, D., Capelletti, M., Yelensky, R., Otto, G., Parker, A., Jarosz, M., Curran, J.A., Balasubramanian, S., Bloom, T., Brennan, K.W., Donahue, A., Downing, S.R., Frampton, G.M., Garcia, L., Juhn, F., Mitchell, K.C., White, E., White, J., Zwirko, Z., Peretz, T., Nechushtan, H., Soussan-Gutman, L., Kim, J., Sasaki, H., Kim, H.R., Park, S.I., Ercan, D., Sheehan, C.E., Ross, J.S., Cronin, M.T., Janne, P.A., Stephens, P.J., 2012. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. *Nature Medicine.* **18,** 382-384.

Lu, H.Y., Wang, X.J., Mao, W.M., 2013. Targeted therapies in small cell lung cancer. *Oncology Letters.* **5,** 3-11.

Ludovini, V., Pistola, L., Gregorc, V., Floriani, I., Rulli, E., Piattoni, S., Di Carlo, L., Semeraro, A., Darwish, S., Tofanetti, F.R., Stocchi, L., Mihaylova, Z., Bellezza, G., Del Sordo, R., Daddi, G., Crino, L., Tonato, M., 2008. Plasma DNA, microsatellite alterations, and p53 tumor mutations are associated with disease-free survival in radically resected non-small cell lung cancer patients: a study of the perugia multidisciplinary team for thoracic oncology. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **3,** 365-373.

Lutz, S., Berk, L., Chang, E., Chow, E., Hahn, C., Hoskin, P., Howell, D., Konski, A., Kachnic, L., Lo, S., Sahgal, A., Silverman, L., von Gunten, C., Mendel, E., Vassil, A., Bruner, D.W., Hartsell, W., American Society for Radiation Oncology (ASTRO), 2011. Palliative radiotherapy for bone metastases: an ASTRO evidence-based guideline. *International Journal of Radiation Oncology, Biology, Physics.* **79,** 965-976.

Madhavan, D., Wallwiener, M., Bents, K., Zucknick, M., Nees, J., Schott, S., Cuk, K., Riethdorf, S., Trumpp, A., Pantel, K., Sohn, C., Schneeweiss, A., Surowy, H., Burwinkel, B., 2014. Plasma DNA integrity as a biomarker for primary and metastatic breast cancer and potential marker for early diagnosis. *Breast Cancer Research and Treatment.* **146,** 163-174.

Maeshima, A.M., Tochigi, N., Yoshida, A., Asamura, H., Tsuta, K., Tsuda, H., 2010. Histological scoring for small lung adenocarcinomas 2 cm or less in diameter: a reliable prognostic indicator. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **5,** 333-339.

Mandel, P. and Métais, P., 1948. Les acides nucléiques du plasma sanguin chez l'homme. **142,** 241-243.

Manser, R.L., Irving, L.B., Stone, C., Byrnes, G., Abramson, M., Campbell, D., 2004. Screening for lung cancer. *Cochrane Database of Systematic Reviews (Online).* **(1),** CD001991.

Mao, L., 2001. Molecular abnormalities in lung carcinogenesis and their potential clinical implications. *Lung Cancer (Amsterdam, Netherlands).* **34 Suppl 2,** S27-34.

Marangoni, E. and Poupon, M.F., 2014. Patient-derived tumour xenografts as models for breast cancer drug development. *Current Opinion in Oncology.* **26,** 556-561.

Marchetti, A., Felicioni, L., Malatesta, S., Grazia Sciarrotta, M., Guetti, L., Chella, A., Viola, P., Pullara, C., Mucilli, F., Buttitta, F., 2011. Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **29,** 3574-3579.

Massion, P.P., Taflan, P.M., Shyr, Y., Rahman, S.M., Yildiz, P., Shakthour, B., Edgerton, M.E., Ninan, M., Andersen, J.J., Gonzalez, A.L., 2004. Early involvement of the phosphatidylinositol 3-kinase/Akt pathway in lung cancer progression. *American Journal of Respiratory and Critical Care Medicine.* **170,** 1088-1094.

Mattioni, M., Chinzari, P., Soddu, S., Strigari, L., Cilenti, V., Mastropasqua, E., 2013. Serum p53 antibody detection in patients with impaired lung function. *BMC Cancer.* **13,** 62-2407-13-62.

McCaughan, F., Pole, J.C., Bankier, A.T., Konfortov, B.A., Carroll, B., Falzon, M., Rabbitts, T.H., George, P.J., Dear, P.H., Rabbitts, P.H., 2010. Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. *American Journal of Respiratory and Critical Care Medicine.* **182,** 83-91.

McClure, D.E., 1999. Clinical pathology and sample collection in the laboratory rodent. *The Veterinary Clinics of North America.Exotic Animal Practice.* **2,** 565-90, vi.

Mechanic, L.E., Marrogi, A.J., Welsh, J.A., Bowman, E.D., Khan, M.A., Enewold, L., Zheng, Y.L., Chanock, S., Shields, P.G., Harris, C.C., 2005. Polymorphisms in XPD and TP53 and mutation in human lung cancer. *Carcinogenesis.* **26,** 597-604.

Melamed, M.R., Flehinger, B.J., Zaman, M.B., Heelan, R.T., Perchick, W.A., Martini, N., 1984. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest.* **86,** 44-53.

Mengual, L., Burset, M., Marin-Aguilera, M., Ribal, M.J., Alcaraz, A., 2008. Multiplex preamplification of specific cDNA targets prior to gene expression analysis by TaqMan Arrays. *BMC Research Notes.* **1,** 21-0500-1-21.

Mercer, K., Giblett, S., Green, S., Lloyd, D., DaRocha Dias, S., Plumb, M., Marais, R., Pritchard, C., 2005. Expression of endogenous oncogenic V600EB-raf induces proliferation and developmental defects in mice and transformation of primary fibroblasts. *Cancer Research.* **65,** 11493-11500.

Mery, C.M., Pappas, A.N., Bueno, R., Colson, Y.L., Linden, P., Sugarbaker, D.J., Jaklitsch, M.T., 2005. Similar long-term survival of elderly patients with non-small cell lung cancer treated with lobectomy or wedge resection within the surveillance, epidemiology, and end results database. *Chest.* **128,** 237-245.

Miyoshi, T., Satoh, Y., Okumura, S., Nakagawa, K., Shirakusa, T., Tsuchiya, E., Ishikawa, Y., 2003. Early-stage lung adenocarcinomas with a micropapillary pattern, a distinct pathologic marker for a significantly poor prognosis. *The American Journal of Surgical Pathology.* **27,** 101-109.

Mogi, A. and Kuwano, H., 2011. TP53 mutations in nonsmall cell lung cancer. *Journal of Biomedicine and Biotechnology.* **2011,** 583929.

Mohan, S., Heitzer, E., Ulz, P., Lafer, I., Lax, S., Auer, M., Pichler, M., Gerger, A., Eisner, F., Hoefler, G., Bauernhofer, T., Geigl, J.B., Speicher, M.R., 2014. Changes in colorectal carcinoma genomes under anti-EGFR therapy identified by whole-genome plasma DNA sequencing. *PLoS Genetics.* **10,** e1004271.

Mok, T., Wu, Y.L., Lee, J.S., Yu, C.J., Sriuranpong, V., Sandoval-Tan, J., Ladrera, G., Thongprasert, S., Srimuninnimit, V., Liao, M., Zhu, Y., Zhou, C., Fuerte, F., Margono, B., Wen, W., Tsai, J., Truman, M., Klughammer, B., Shames, D.S., Wu, L., 2015. Detection and Dynamic Changes of EGFR Mutations from Circulating Tumor DNA as a Predictor of Survival Outcomes in NSCLC Patients Treated with First-line Intercalated Erlotinib and Chemotherapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **21,** 3196-3203.

Molina, J.R., Yang, P., Cassivi, S.D., Schild, S.E., Adjei, A.A., 2008. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clinic Proceedings.* **83,** 584-594.

Mori, M., Rao, S.K., Popper, H.H., Cagle, P.T., Fraire, A.E., 2001. Atypical adenomatous hyperplasia of the lung: a probable forerunner in the development of adenocarcinoma of the lung. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.* **14,** 72-84.

Mori, N., Yokota, J., Akiyama, T., Sameshima, Y., Okamoto, A., Mizoguchi, H., Toyoshima, K., Sugimura, T., Terada, M., 1990. Variable mutations of the RB gene in small-cell lung carcinoma. *Oncogene.* **5,** 1713-1717.

Moro, D., Brichon, P.Y., Brambilla, E., Veale, D., Labat, F., Brambilla, C., 1994. Basaloid bronchial carcinoma. A histologic group with a poor prognosis. *Cancer.* **73,** 2734-2739.

Mouliere, F., El Messaoudi, S., Gongora, C., Guedj, A.S., Robert, B., Del Rio, M., Molina, F., Lamy, P.J., Lopez-Crapez, E., Mathonnet, M., Ychou, M., Pezet, D., Thierry, A.R., 2013. Circulating Cell-Free DNA from Colorectal Cancer Patients May Reveal High KRAS or BRAF Mutation Load. *Translational Oncology.* **6,** 319-328.

Mouliere, F., Robert, B., Arnau Peyrotte, E., Del Rio, M., Ychou, M., Molina, F., Gongora, C., Thierry, A.R., 2011. High fragmentation characterizes tumour-derived circulating DNA. *PloS One.* **6,** e23418.

Murray, J.F., 2010. The structure and function of the lung. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union Against Tuberculosis and Lung Disease.* **14,** 391-396.

Murtaza, M., Dawson, S.J., Tsui, D.W., Gale, D., Forshew, T., Piskorz, A.M., Parkinson, C., Chin, S.F., Kingsbury, Z., Wong, A.S., Marass, F., Humphray, S., Hadfield, J., Bentley, D., Chin, T.M., Brenton, J.D., Caldas, C., Rosenfeld, N., 2013. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature.* **497,** 108-112.

Namati, E., Thiesse, J., Sieren, J.C., Ross, A., Hoffman, E.A., McLennan, G., 2010a. Longitudinal assessment of lung cancer progression in the mouse using in vivo micro-CT imaging. *Medical Physics.* **37,** 4793-4805.

Namati, E., Thiesse, J., Sieren, J.C., Ross, A., Hoffman, E.A., McLennan, G., 2010b. Longitudinal assessment of lung cancer progression in the mouse using in vivo micro-CT imaging. *Medical Physics.* **37,** 4793-4805.

National Lung Screening Trial Research Team, Aberle, D.R., Adams, A.M., Berg, C.D., Black, W.C., Clapp, J.D., Fagerstrom, R.M., Gareen, I.F., Gatsonis, C., Marcus, P.M., Sicks, J.D., 2011a. Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England Journal of Medicine.* **365,** 395-409.

National Lung Screening Trial Research Team, Aberle, D.R., Berg, C.D., Black, W.C., Church, T.R., Fagerstrom, R.M., Galen, B., Gareen, I.F., Gatsonis, C., Goldin, J., Gohagan, J.K., Hillman, B., Jaffe, C., Kramer, B.S., Lynch, D., Marcus, P.M., Schnall, M., Sullivan, D.C., Sullivan, D., Zylak, C.J., 2011b. The

National Lung Screening Trial: overview and study design. *Radiology.* **258,** 243-253.

Newman, A.M., Bratman, S.V., To, J., Wynne, J.F., Eclov, N.C., Modlin, L.A., Liu, C.L., Neal, J.W., Wakelee, H.A., Merritt, R.E., Shrager, J.B., Loo, B.W.,Jr, Alizadeh, A.A., Diehn, M., 2014. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature Medicine.* **20,** 548-554.

Nie, K., Jia, Y., Zhang, X., 2015. Cell-free circulating tumor DNA in plasma/serum of non-small cell lung cancer. *Tumour Biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine.* **36,** 7-19.

Noutsias, M., Rohde, M., Block, A., Klippert, K., Lettau, O., Blunert, K., Hummel, M., Kuhl, U., Lehmkuhl, H., Hetzer, R., Rauch, U., Poller, W., Pauschinger, M., Schultheiss, H.P., Volk, H.D., Kotsch, K., 2008. Preamplification techniques for real-time RT-PCR analyses of endomyocardial biopsies. *BMC Molecular Biology.* **9,** 3-2199-9-3.

NSCLC Meta-Analyses Collaborative Group, 2008. Chemotherapy in addition to supportive care improves survival in advanced non-small-cell lung cancer: a systematic review and meta-analysis of individual patient data from 16 randomized controlled trials. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **26,** 4617-4625.

NSCLC Meta-analyses Collaborative Group, Arriagada, R., Auperin, A., Burdett, S., Higgins, J.P., Johnson, D.H., Le Chevalier, T., Le Pechoux, C., Parmar, M.K., Pignon, J.P., Souhami, R.L., Stephens, R.J., Stewart, L.A., Tierney, J.F., Tribodet, H., van Meerbeeck, J., 2010. Adjuvant chemotherapy, with or without postoperative radiotherapy, in operable non-small-cell lung cancer: two meta-analyses of individual patient data. *Lancet.* **375,** 1267-1277.

Nygaard, A.D., Garm Spindler, K.L., Pallisgaard, N., Andersen, R.F., Jakobsen, A., 2013. The prognostic value of KRAS mutated plasma DNA in advanced non-small cell lung cancer. *Lung Cancer (Amsterdam, Netherlands).* **79,** 312-317.

Nygaard, A.D., Holdgaard, P.C., Spindler, K.L., Pallisgaard, N., Jakobsen, A., 2014. The correlation between cell-free DNA and tumour burden was estimated by PET/CT in patients with advanced NSCLC. *British Journal of Cancer.* **110,** 363-368.

Office for National Statistics, 2011. Cancer Survival in England: Patients Diagnosed 2005-2009 and Followed up to 2010 .

Ohashi, K., Sequist, L.V., Arcila, M.E., Moran, T., Chmielecki, J., Lin, Y.L., Pan, Y., Wang, L., de Stanchina, E., Shien, K., Aoe, K., Toyooka, S., Kiura, K.,

Fernandez-Cuesta, L., Fidias, P., Yang, J.C., Miller, V.A., Riely, G.J., Kris, M.G., Engelman, J.A., Vnencak-Jones, C.L., Dias-Santagata, D., Ladanyi, M., Pao, W., 2012. Lung cancers with acquired resistance to EGFR inhibitors occasionally harbor BRAF gene mutations but lack mutations in KRAS, NRAS, or MEK1. *Proceedings of the National Academy of Sciences of the United States of America.* **109,** E2127-33.

Oken, M.M., Hocking, W.G., Kvale, P.A., Andriole, G.L., Buys, S.S., Church, T.R., Crawford, E.D., Fouad, M.N., Isaacs, C., Reding, D.J., Weissfeld, J.L., Yokochi, L.A., O'Brien, B., Ragard, L.R., Rathmell, J.M., Riley, T.L., Wright, P., Caparaso, N., Hu, P., Izmirlian, G., Pinsky, P.F., Prorok, P.C., Kramer, B.S., Miller, A.B., Gohagan, J.K., Berg, C.D., PLCO Project Team, 2011. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA : The Journal of the American Medical Association.* **306,** 1865-1873.

Okuda, K., Sasaki, H., Yukiue, H., Yano, M., Fujii, Y., 2008. Met gene copy number predicts the prognosis for completely resected non-small cell lung cancer. *Cancer Science.* **99,** 2280-2285.

Okudela, K., Suzuki, M., Kageyama, S., Bunai, T., Nagura, K., Igarashi, H., Takamochi, K., Suzuki, K., Yamada, T., Niwa, H., Ohashi, R., Ogawa, H., Mori, H., Kitamura, H., Kaneko, T., Tsuneyoshi, T., Sugimura, H., 2007. PIK3CA mutation and amplification in human lung cancer. *Pathology International.* **57,** 664-671.

Otterson, G., Lin, A., Kay, F., 1992. Genetic etiology of lung cancer. *Oncology (Williston Park, N.Y.).* **6,** 97-104, 107; discussion 108, 111-2.

Owczarzy, R., You, Y., Groth, C.L., Tataurov, A.V., 2011. Stability and mismatch discrimination of locked nucleic acid-DNA duplexes. *Biochemistry.* **50,** 9352-9367.

Paci, M., Maramotti, S., Bellesia, E., Formisano, D., Albertazzi, L., Ricchetti, T., Ferrari, G., Annessi, V., Lasagni, D., Carbonelli, C., De Franco, S., Brini, M., Sgarbi, G., Lodi, R., 2009. Circulating plasma DNA as diagnostic biomarker in non-small cell lung cancer. *Lung Cancer (Amsterdam, Netherlands).* **64,** 92-97.

Page, K., Hava, N., Ward, B., Brown, J., Guttery, D.S., Ruangpratheep, C., Blighe, K., Sharma, A., Walker, R.A., Coombes, R.C., Shaw, J.A., 2011. Detection of HER2 amplification in circulating free DNA in patients with breast cancer. *British Journal of Cancer.* **104,** 1342-1348.

Page, K., Guttery, D., Fernandez-Garcia, D., Hills, A., Hastings, R., Luo, J., Goddard, K., Shahin, V., Woodley-Barker, L., Rosales, B., Coombes, RC., Stebbing, J., Shaw, JA., 2016. Next generation sequencing of circulating cell-

free DNA evaluating mutations and gene amplification in metastatic breast cancer. *Clinical Chemistry*. [in press]

Park, K.S., Liang, M.C., Raiser, D.M., Zamponi, R., Roach, R.R., Curtis, S.J., Walton, Z., Schaffer, B.E., Roake, C.M., Zmoos, A.F., Kriegel, C., Wong, K.K., Sage, J., Kim, C.F., 2011. Characterization of the cell of origin for small cell lung cancer. *Cell Cycle (Georgetown, Tex.)*. **10,** 2806-2815.

Pastorino, U., Bellomi, M., Landoni, C., De Fiori, E., Arnaldi, P., Picchio, M., Pelosi, G., Boyle, P., Fazio, F., 2003. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet*. **362,** 593-597.

Paz-Ares, L., Soulieres, D., Melezinek, I., Moecks, J., Keil, L., Mok, T., Rosell, R., Klughammer, B., 2010. Clinical outcomes in non-small-cell lung cancer patients with EGFR mutations: pooled analysis. *Journal of Cellular and Molecular Medicine*. **14,** 51-69.

Peifer, M., Fernandez-Cuesta, L., Sos, M.L., George, J., Seidel, D., Kasper, L.H., Plenker, D., Leenders, F., Sun, R., Zander, T., Menon, R., Koker, M., Dahmen, I., Muller, C., Di Cerbo, V., Schildhaus, H.U., Altmuller, J., Baessmann, I., Becker, C., de Wilde, B., Vandesompele, J., Bohm, D., Ansen, S., Gabler, F., Wilkening, I., Heynck, S., Heuckmann, J.M., Lu, X., Carter, S.L., Cibulskis, K., Banerji, S., Getz, G., Park, K.S., Rauh, D., Grutter, C., Fischer, M., Pasqualucci, L., Wright, G., Wainer, Z., Russell, P., Petersen, I., Chen, Y., Stoelben, E., Ludwig, C., Schnabel, P., Hoffmann, H., Muley, T., Brockmann, M., Engel-Riedel, W., Muscarella, L.A., Fazio, V.M., Groen, H., Timens, W., Sietsma, H., Thunnissen, E., Smit, E., Heideman, D.A., Snijders, P.J., Cappuzzo, F., Ligorio, C., Damiani, S., Field, J., Solberg, S., Brustugun, O.T., Lund-Iversen, M., Sanger, J., Clement, J.H., Soltermann, A., Moch, H., Weder, W., Solomon, B., Soria, J.C., Validire, P., Besse, B., Brambilla, E., Brambilla, C., Lantuejoul, S., Lorimier, P., Schneider, P.M., Hallek, M., Pao, W., Meyerson, M., Sage, J., Shendure, J., Schneider, R., Buttner, R., Wolf, J., Nurnberg, P., Perner, S., Heukamp, L.C., Brindle, P.K., Haas, S., Thomas, R.K., 2012. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nature Genetics*. **44,** 1104-1110.

Pellegrini, C., Falleni, M., Marchetti, A., Cassani, B., Miozzo, M., Buttitta, F., Roncalli, M., Coggi, G., Bosari, S., 2003. HER-2/Neu alterations in non-small cell lung cancer: a comprehensive evaluation by real time reverse transcription-PCR, fluorescence in situ hybridization, and immunohistochemistry. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*. **9,** 3645-3652.

Pignon, J.P., Tribodet, H., Scagliotti, G.V., Douillard, J.Y., Shepherd, F.A., Stephens, R.J., Dunant, A., Torri, V., Rosell, R., Seymour, L., Spiro, S.G., Rolland, E., Fossati, R., Aubert, D., Ding, K., Waller, D., Le Chevalier, T., LACE

Collaborative Group, 2008. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **26,** 3552-3559.

Piotrowska, Z., Niederst, M.J., Karlovich, C.A., Wakelee, H.A., Neal, J.W., Mino-Kenudson, M., Fulton, L., Hata, A.N., Lockerman, E.L., Kalsy, A., Digumarthy, S., Muzikansky, A., Raponi, M., Garcia, A.R., Mulvey, H.E., Parks, M.K., DiCecca, R.H., Dias-Santagata, D., Iafrate, A.J., Shaw, A.T., Allen, A.R., Engelman, J.A., Sequist, L.V., 2015. Heterogeneity Underlies the Emergence of EGFRT790 Wild-Type Clones Following Treatment of T790M-Positive Cancers with a Third-Generation EGFR Inhibitor. *Cancer Discovery.* **5,** 713-722.

Pitteri, S.J., Faca, V.M., Kelly-Spratt, K.S., Kasarda, A.E., Wang, H., Zhang, Q., Newcomb, L., Krasnoselsky, A., Paczesny, S., Choi, G., Fitzgibbon, M., McIntosh, M.W., Kemp, C.J., Hanash, S.M., 2008. Plasma proteome profiling of a mouse model of breast cancer identifies a set of up-regulated proteins in common with human breast cancer cells. *Journal of Proteome Research.* **7,** 1481-1489.

Ponticiello, A., Barra, E., Giani, U., Bocchino, M., Sanduzzi, A., 2000. P53 immunohistochemistry can identify bronchial dysplastic lesions proceeding to lung cancer: a prospective study. *The European Respiratory Journal.* **15,** 547-552.

Prophylactic Cranial Irradiation Overview Collaborative Group, 2000. Cranial irradiation for preventing brain metastases of small cell lung cancer in patients in complete remission. *The Cochrane Database of Systematic Reviews.* **(4),** CD002805.

Pujol, J.L., Carestia, L., Daures, J.P., 2000. Is there a case for cisplatin in the treatment of small-cell lung cancer? A meta-analysis of randomized trials of a cisplatin-containing regimen versus a regimen without this alkylating agent. *British Journal of Cancer.* **83,** 8-15.

Qiu, T., Guo, H., Zhao, H., Wang, L., Zhang, Z., 2015. Next-generation sequencing for molecular diagnosis of lung adenocarcinoma specimens obtained by fine needle aspiration cytology. *Scientific Reports.* **5,** 11317.

Rago, C., Huso, D.L., Diehl, F., Karim, B., Liu, G., Papadopoulos, N., Samuels, Y., Velculescu, V.E., Vogelstein, B., Kinzler, K.W., Diaz, L.A.,Jr, 2007. Serial assessment of human tumor burdens in mice by the analysis of circulating DNA. *Cancer Research.* **67,** 9364-9370.

Rainer, T.H., Wong, L.K., Lam, W., Yuen, E., Lam, N.Y., Metreweli, C., Lo, Y.M., 2003. Prognostic use of circulating plasma nucleic acid concentrations in patients with acute stroke. *Clinical Chemistry.* **49,** 562-569.

Ramakers, C., Ruijter, J.M., Deprez, R.H., Moorman, A.F., 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters.* **339,** 62-66.

Ramalingam, S., Goss, G., Rosell, R., Schmid-Bindert, G., Zaric, B., Andric, Z., Bondarenko, I., Komov, D., Ceric, T., Khuri, F., Samarzija, M., Felip, E., Ciuleanu, T., Hirsh, V., Wehler, T., Spicer, J., Salgia, R., Shapiro, G., Sheldon, E., Teofilovici, F., Vukovic, V., Fennell, D., 2015. A randomized phase II study of ganetespib, a heat shock protein 90 inhibitor, in combination with docetaxel in second-line therapy of advanced non-small cell lung cancer (GALAXY-1). *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO.* **26,** 1741-1748.

Rhodes, A., Wort, S.J., Thomas, H., Collinson, P., Bennett, E.D., 2006. Plasma DNA concentration as a predictor of mortality and sepsis in critically ill patients. *Critical Care (London, England).* **10,** R60.

Richmond, A. and Su, Y., 2008. Mouse xenograft models vs GEM models for human cancer therapeutics. *Disease Models and Mechanisms.* **1,** 78-82.

Ried, T., Petersen, I., Holtgreve-Grez, H., Speicher, M.R., Schrock, E., du Manoir, S., Cremer, T., 1994. Mapping of multiple DNA gains and losses in primary small cell lung carcinomas by comparative genomic hybridization. *Cancer Research.* **54,** 1801-1806.

Rimkunas, V.M., Crosby, K.E., Li, D., Hu, Y., Kelly, M.E., Gu, T.L., Mack, J.S., Silver, M.R., Zhou, X., Haack, H., 2012. Analysis of receptor tyrosine kinase ROS1-positive tumors in non-small cell lung cancer: identification of a FIG-ROS1 fusion. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **18,** 4449-4457.

Robinson, L.A., Ruckdeschel, J.C., Wagner, H.,Jr, Stevens, C.W., American College of Chest Physicians, 2007. Treatment of non-small cell lung cancer-stage IIIA: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest.* **132,** 243S-265S.

Rodriguez, E., Mannion, L., D'Santos, P., Griffiths, M., Arends, M.J., Brindle, K.M., Lyons, S.K., 2014. Versatile and enhanced tumour modelling in mice via somatic cell transduction. *The Journal of Pathology.* **232,** 449-457.

Rooney, C., Geh, C., Williams, V., Heuckmann, J.M., Menon, R., Schneider, P., Al-Kadhimi, K., Dymond, M., Smith, N.R., Baker, D., French, T., Smith, P.D., Harrington, E.A., Barrett, J.C., Kilgour, E., 2016. Characterization of FGFR1 Locus in sqNSCLC Reveals a Broad and Heterogeneous Amplicon. *PloS One.* **11,** e0149628.

Ross, J.S., Wang, K., Elkadi, O.R., Tarasen, A., Foulke, L., Sheehan, C.E., Otto, G.A., Palmer, G., Yelensky, R., Lipson, D., Chmielecki, J., Ali, S.M., Elvin, J., Morosini, D., Miller, V.A., Stephens, P.J., 2014. Next-generation sequencing reveals frequent consistent genomic alterations in small cell undifferentiated lung cancer. *Journal of Clinical Pathology.* **67,** 772-776.

Roth, C., Kasimir-Bauer, S., Pantel, K., Schwarzenbach, H., 2011. Screening for circulating nucleic acids and caspase activity in the peripheral blood as potential diagnostic tools in lung cancer. *Molecular Oncology.* **5,** 281-291.

Rowell, N.P. and Williams, C.J., 2001. Radical radiotherapy for stage I/II non-small cell lung cancer in patients not sufficiently fit for or declining surgery (medically inoperable): a systematic review. *Thorax.* **56,** 628-638.

Rudin, C.M., Durinck, S., Stawiski, E.W., Poirier, J.T., Modrusan, Z., Shames, D.S., Bergbower, E.A., Guan, Y., Shin, J., Guillory, J., Rivers, C.S., Foo, C.K., Bhatt, D., Stinson, J., Gnad, F., Haverty, P.M., Gentleman, R., Chaudhuri, S., Janakiraman, V., Jaiswal, B.S., Parikh, C., Yuan, W., Zhang, Z., Koeppen, H., Wu, T.D., Stern, H.M., Yauch, R.L., Huffman, K.E., Paskulin, D.D., Illei, P.B., Varella-Garcia, M., Gazdar, A.F., de Sauvage, F.J., Bourgon, R., Minna, J.D., Brock, M.V., Seshagiri, S., 2012. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nature Genetics.* **44,** 1111-1116.

Rumore, P.M. and Steinman, C.R., 1990. Endogenous circulating DNA in systemic lupus erythematosus. Occurrence as multimeric complexes bound to histone. *The Journal of Clinical Investigation.* **86,** 69-74.

Sakamoto, H., Shimizu, J., Horio, Y., Ueda, R., Takahashi, T., Mitsudomi, T., Yatabe, Y., 2007. Disproportionate representation of KRAS gene mutation in atypical adenomatous hyperplasia, but even distribution of EGFR gene mutation from preinvasive to invasive adenocarcinomas. *The Journal of Pathology.* **212,** 287-294.

Sanders, R., Huggett, J.F., Bushell, C.A., Cowen, S., Scott, D.J., Foy, C.A., 2011. Evaluation of digital PCR for absolute DNA quantification. *Analytical Chemistry.* **83,** 6474-6484.

Schneider, B.J., 2008. Management of recurrent small cell lung cancer. *Journal of the National Comprehensive Cancer Network : JNCCN.* **6,** 323-331.

Schneider, F., Luvison, A., Cieply, K., Dacic, S., 2013. Sex-determining region Y-box 2 amplification in preneoplastic squamous lesions of the lung. *Human Pathology.* **44,** 706-711.

Schultheis, A.M., Bos, M., Schmitz, K., Wilsberg, L., Binot, E., Wolf, J., Buttner, R., Schildhaus, H.U., 2013. Fibroblast growth factor receptor 1 (FGFR1)

amplification is a potential therapeutic target in small-cell lung cancer. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.*

Schwarzenbach, H., Stoehlmacher, J., Pantel, K., Goekkurt, E., 2008. Detection and monitoring of cell-free DNA in blood of patients with colorectal cancer. *Annals of the New York Academy of Sciences.* **1137,** 190-196.

Scott, W.J., Howington, J., Feigenberg, S., Movsas, B., Pisters, K., American College of Chest Physicians, 2007. Treatment of non-small cell lung cancer stage I and stage II: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest.* **132,** 234S-242S.

Semenova, E.A., Nagel, R., Berns, A., 2015. Origins, genetic landscape, and emerging therapies of small cell lung cancer. *Genes and Development.* **29,** 1447-1462.

Seo, J.S., Ju, Y.S., Lee, W.C., Shin, J.Y., Lee, J.K., Bleazard, T., Lee, J., Jung, Y.J., Kim, J.O., Shin, J.Y., Yu, S.B., Kim, J., Lee, E.R., Kang, C.H., Park, I.K., Rhee, H., Lee, S.H., Kim, J.I., Kang, J.H., Kim, Y.T., 2012. The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Research.* **22,** 2109-2119.

Shaikh, A.C. and Sadowski, P.D., 1997. The Cre recombinase cleaves the lox site in trans. *The Journal of Biological Chemistry.* **272,** 5695-5702.

Shao, D., Lin, Y., Liu, J., Wan, L., Liu, Z., Cheng, S., Fei, L., Deng, R., Wang, J., Chen, X., Liu, L., Gu, X., Liang, W., He, P., Wang, J., Ye, M., He, J., 2016. A targeted next-generation sequencing method for identifying clinically relevant mutation profiles in lung adenocarcinoma. *Scientific Reports.* **6,** 22338.

Shaw, J.A., Page, K., Blighe, K., Hava, N., Guttery, D., Ward, B., Brown, J., Ruangpratheep, C., Stebbing, J., Payne, R., Palmieri, C., Cleator, S., Walker, R.A., Coombes, R.C., 2012. Genomic analysis of circulating cell-free DNA infers breast cancer dormancy. *Genome Research.* **22,** 220-231.

Shigematsu, H., Lin, L., Takahashi, T., Nomura, M., Suzuki, M., Wistuba, I.I., Fong, K.M., Lee, H., Toyooka, S., Shimizu, N., Fujisawa, T., Feng, Z., Roth, J.A., Herz, J., Minna, J.D., Gazdar, A.F., 2005. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *Journal of the National Cancer Institute.* **97,** 339-346.

Shinozaki, M., O'Day, S.J., Kitago, M., Amersi, F., Kuo, C., Kim, J., Wang, H.J., Hoon, D.S., 2007. Utility of circulating B-RAF DNA mutation in serum for monitoring melanoma patients receiving biochemotherapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **13,** 2068-2074.

Shiraishi, M., Noguchi, M., Shimosato, Y., Sekiya, T., 1989. Amplification of protooncogenes in surgical specimens of human lung carcinomas. *Cancer Research.* **49,** 6474-6479.

Shivapurkar, N., Virmani, A.K., Wistuba, I.I., Milchgrub, S., Mackay, B., Minna, J.D., Gazdar, A.F., 1999. Deletions of chromosome 4 at multiple sites are frequent in malignant mesothelioma and small cell lung carcinoma. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **5,** 17-23.

Sholl, L.M., Yeap, B.Y., Iafrate, A.J., Holmes-Tisch, A.J., Chou, Y.P., Wu, M.T., Goan, Y.G., Su, L., Benedettini, E., Yu, J., Loda, M., Janne, P.A., Christiani, D.C., Chirieac, L.R., 2009. Lung adenocarcinoma with EGFR amplification has distinct clinicopathologic and molecular features in never-smokers. *Cancer Research.* **69,** 8341-8348.

Singh, B., Reddy, P.G., Goberdhan, A., Walsh, C., Dao, S., Ngai, I., Chou, T.C., O-Charoenrat, P., Levine, A.J., Rao, P.H., Stoffel, A., 2002. p53 regulates cell survival by inhibiting PIK3CA in squamous cell carcinomas. *Genes and Development.* **16,** 984-993.

Sirera, R., Bremnes, R.M., Cabrera, A., Jantus-Lewintre, E., Sanmartin, E., Blasco, A., Del Pozo, N., Rosell, R., Guijarro, R., Galbis, J., Sanchez, J.J., Camps, C., 2011. Circulating DNA is a useful prognostic factor in patients with advanced non-small cell lung cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **6,** 286-290.

Slotman, B.J., Faivre-Finn, C., Kramer, G.W., Rankin, E., Snee, M., Hatton, M., Postmus, P.E., Collette, L., Musat, E., Senan, S., 2008. Prophylactic cranial irradiation in patients with extensive disease caused by small-cell lung cancer responsive to chemotherapy: fewer symptomatic brain metastases and improved survival. *Nederlands Tijdschrift Voor Geneeskunde.* **152,** 1000-1004.

Sobue, T., Moriyama, N., Kaneko, M., Kusumoto, M., Kobayashi, T., Tsuchiya, R., Kakinuma, R., Ohmatsu, H., Nagai, K., Nishiyama, H., Matsui, E., Eguchi, K., 2002. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **20,** 911-920.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y., Mano, H., 2007. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature.* **448,** 561-566.

Sone, S., Takashima, S., Li, F., Yang, Z., Honda, T., Maruyama, Y., Hasegawa, M., Yamanda, T., Kubo, K., Hanamura, K., Asakura, K., 1998. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet.* **351,** 1242-1245.

Sone, T., Kasahara, K., Kimura, H., Nishio, K., Mizuguchi, M., Nakatsumi, Y., Shibata, K., Waseda, Y., Fujimura, M., Nakao, S., 2007. Comparative analysis of epidermal growth factor receptor mutations and gene amplification as predictors of gefitinib efficacy in Japanese patients with nonsmall cell lung cancer. *Cancer.* **109,** 1836-1844.

Sorenson, G.D., Pribish, D.M., Valone, F.H., Memoli, V.A., Bzik, D.J., Yao, S.L., 1994. Soluble normal and mutated DNA sequences from single-copy genes in human blood. *Cancer Epidemiology, Biomarkers and Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology.* **3,** 67-71.

Sos, M.L., Dietlein, F., Peifer, M., Schottle, J., Balke-Want, H., Muller, C., Koker, M., Richters, A., Heynck, S., Malchers, F., Heuckmann, J.M., Seidel, D., Eyers, P.A., Ullrich, R.T., Antonchick, A.P., Vintonyak, V.V., Schneider, P.M., Ninomiya, T., Waldmann, H., Buttner, R., Rauh, D., Heukamp, L.C., Thomas, R.K., 2012. A framework for identification of actionable cancer genome dependencies in small cell lung cancer. *Proceedings of the National Academy of Sciences of the United States of America.* **109,** 17034-17039.

Sozzi, G., Conte, D., Leon, M., Ciricione, R., Roz, L., Ratcliffe, C., Roz, E., Cirenei, N., Bellomi, M., Pelosi, G., Pierotti, M.A., Pastorino, U., 2003. Quantification of free circulating DNA as a diagnostic marker in lung cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **21,** 3902-3908.

Sozzi, G., Conte, D., Mariani, L., Lo Vullo, S., Roz, L., Lombardo, C., Pierotti, M.A., Tavecchio, L., 2001. Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. *Cancer Research.* **61,** 4675-4678.

Sozzi, G., Musso, K., Ratcliffe, C., Goldstraw, P., Pierotti, M.A., Pastorino, U., 1999. Detection of microsatellite alterations in plasma DNA of non-small cell lung cancer patients: a prospect for early diagnosis. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **5,** 2689-2692.

Sozzi, G., Veronese, M.L., Negrini, M., Baffa, R., Cotticelli, M.G., Inoue, H., Tornielli, S., Pilotti, S., De Gregorio, L., Pastorino, U., Pierotti, M.A., Ohta, M., Huebner, K., Croce, C.M., 1996. The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell.* **85,** 17-26.

Spigel, D.R., 2012. Treatment update in small-cell lung cancer: from limited to extensive disease. *Current Treatment Options in Oncology.* **13,** 505-515.

Staaf, J., Isaksson, S., Karlsson, A., Jonsson, M., Johansson, L., Jonsson, P., Botling, J., Micke, P., Baldetorp, B., Planck, M., 2013. Landscape of somatic allelic imbalances and copy number alterations in human lung carcinoma. *International Journal of Cancer.Journal International Du Cancer.* **132,** 2020-2031.

Stahel, R., Ginsberg, R., Havemann, K., 1989. Staging and prognostic factors in small cell lung cancer: a consensus report. *Lung Cancer.* **5,** 119–126.

Stewart, E.L., Tan, S.Z., Liu, G., Tsao, M.S., 2015. Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations-a review. *Translational Lung Cancer Research.* **4,** 67-81.

Sugio, K., Kishimoto, Y., Virmani, A.K., Hung, J.Y., Gazdar, A.F., 1994. K-ras mutations are a relatively late event in the pathogenesis of lung carcinomas. *Cancer Research.* **54,** 5811-5815.

Summermann, E., Huwer, H., Seitz, G., 1990. Carcinosarcoma of the lung, a tumour which has a poor prognosis and is extremely rarely diagnosed preoperatively. *The Thoracic and Cardiovascular Surgeon.* **38,** 247-250.

Sun, S., Schiller, J.H., Gazdar, A.F., 2007. Lung cancer in never smokers--a different disease. *Nature Reviews.Cancer.* **7,** 778-790.

Sundstrom, S., Bremnes, R.M., Kaasa, S., Aasebo, U., Hatlevoll, R., Dahle, R., Boye, N., Wang, M., Vigander, T., Vilsvik, J., Skovlund, E., Hannisdal, E., Aamdal, S., Norwegian Lung Cancer Study Group, 2002. Cisplatin and etoposide regimen is superior to cyclophosphamide, epirubicin, and vincristine regimen in small-cell lung cancer: results from a randomized phase III trial with 5 years' follow-up. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **20,** 4665-4672.

Suo, Z., Daehli, K.U., Lindboe, C.F., Borgen, E., Bassarova, A., Nesland, J.M., 2004. Real-time PCR quantification of c-erbB-2 gene is an alternative for FISH in the clinical management of breast carcinoma patients. *International Journal of Surgical Pathology.* **12,** 311-318.

Sutherland, K.D., Proost, N., Brouns, I., Adriaensen, D., Song, J.Y., Berns, A., 2011. Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. *Cancer Cell.* **19,** 754-764.

Sutherland, K.D., Song, J.Y., Kwon, M.C., Proost, N., Zevenhoven, J., Berns, A., 2014. Multiple cells-of-origin of mutant K-Ras-induced mouse lung

adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America.* **111,** 4952-4957.

Swensen, S.J., Jett, J.R., Hartman, T.E., Midthun, D.E., Mandrekar, S.J., Hillman, S.L., Sykes, A.M., Aughenbaugh, G.L., Bungum, A.O., Allen, K.L., 2005. CT screening for lung cancer: five-year prospective experience. *Radiology.* **235,** 259-265.

Swystun, L.L., Mukherjee, S., Liaw, P.C., 2011. Breast cancer chemotherapy induces the release of cell-free DNA, a novel procoagulant stimulus. *Journal of Thrombosis and Haemostasis : JTH.* **9,** 2313-2321.

Szpechcinski, A., Chorostowska-Wynimko, J., Struniawski, R., Kupis, W., Rudzinski, P., Langfort, R., Puscinska, E., Bielen, P., Sliwinski, P., Orlowski, T., 2015. Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. *British Journal of Cancer.* **113,** 476-483.

Szpechcinski, A., Rudzinski, P., Kupis, W., Langfort, R., Orlowski, T., Chorostowska-Wynimko, J., 2016. Plasma cell-free DNA levels and integrity in patients with chest radiological findings: NSCLC versus benign lung nodules. *Cancer Letters.* **374,** 202-207.

Taguchi, A., Politi, K., Pitteri, S.J., Lockwood, W.W., Faca, V.M., Kelly-Spratt, K., Wong, C.H., Zhang, Q., Chin, A., Park, K.S., Goodman, G., Gazdar, A.F., Sage, J., Dinulescu, D.M., Kucherlapati, R., Depinho, R.A., Kemp, C.J., Varmus, H.E., Hanash, S.M., 2011. Lung cancer signatures in plasma based on proteome profiling of mouse tumor models. *Cancer Cell.* **20,** 289-299.

Takei, H., Asamura, H., Maeshima, A., Suzuki, K., Kondo, H., Niki, T., Yamada, T., Tsuchiya, R., Matsuno, Y., 2002. Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases. *The Journal of Thoracic and Cardiovascular Surgery.* **124,** 285-292.

Takeuchi, K., Soda, M., Togashi, Y., Suzuki, R., Sakata, S., Hatano, S., Asaka, R., Hamanaka, W., Ninomiya, H., Uehara, H., Lim Choi, Y., Satoh, Y., Okumura, S., Nakagawa, K., Mano, H., Ishikawa, Y., 2012. RET, ROS1 and ALK fusions in lung cancer. *Nature Medicine.* **18,** 378-381.

Takezawa, K., Pirazzoli, V., Arcila, M.E., Nebhan, C.A., Song, X., de Stanchina, E., Ohashi, K., Janjigian, Y.Y., Spitzler, P.J., Melnick, M.A., Riely, G.J., Kris, M.G., Miller, V.A., Ladanyi, M., Politi, K., Pao, W., 2012. HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFRT790M mutation. *Cancer Discovery.* **2,** 922-933.

Tang, X., Shigematsu, H., Bekele, B.N., Roth, J.A., Minna, J.D., Hong, W.K., Gazdar, A.F., Wistuba, I.I., 2005. EGFR tyrosine kinase domain mutations are

detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Research.* **65,** 7568-7572.

Tao, L., van Bragt, M.P., Laudadio, E., Li, Z., 2014. Lineage tracing of mammary epithelial cells using cell-type-specific cre-expressing adenoviruses. *Stem Cell Reports.* **2,** 770-779.

Thierry, A.R., Mouliere, F., Gongora, C., Ollier, J., Robert, B., Ychou, M., Del Rio, M., Molina, F., 2010. Origin and quantification of circulating DNA in mice with human colorectal cancer xenografts. *Nucleic Acids Research.* **38,** 6159-6175.

Thomas, A., Lee, J.H., Abdullaev, Z., Park, K.S., Pineda, M., Saidkhodjaeva, L., Miettinen, M., Wang, Y., Pack, S.D., Giaccone, G., 2014. Characterization of fibroblast growth factor receptor 1 in small-cell lung cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **9,** 567-571.

Thomson, J.G., Rucker, E.B.,3rd, Piedrahita, J.A., 2003. Mutational analysis of loxP sites for efficient Cre-mediated insertion into genomic DNA. *Genesis (New York, N.Y.: 2000).* **36,** 162-167.

Thunnissen, E., Bubendorf, L., Dietel, M., Elmberger, G., Kerr, K., Lopez-Rios, F., Moch, H., Olszewski, W., Pauwels, P., Penault-Llorca, F., Rossi, G., 2012. EML4-ALK testing in non-small cell carcinomas of the lung: a review with recommendations. *Virchows Archiv : An International Journal of Pathology.* **461,** 245-257.

To, E.W., Chan, K.C., Leung, S.F., Chan, L.Y., To, K.F., Chan, A.T., Johnson, P.J., Lo, Y.M., 2003. 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.* **9,** 3254-3259.

Tokumo, M., Toyooka, S., Kiura, K., Shigematsu, H., Tomii, K., Aoe, M., Ichimura, K., Tsuda, T., Yano, M., Tsukuda, K., Tabata, M., Ueoka, H., Tanimoto, M., Date, H., Gazdar, A.F., Shimizu, N., 2005. The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **11,** 1167-1173.

Tominaga, M., Sueoka, N., Irie, K., Iwanaga, K., Tokunaga, O., Hayashi, S., Nakachi, K., Sueoka, E., 2003. Detection and discrimination of preneoplastic and early stages of lung adenocarcinoma using hnRNP B1 combined with the cell cycle-related markers p16, cyclin D1, and Ki-67. *Lung Cancer (Amsterdam, Netherlands).* **40,** 45-53.

Toyooka, S., Tsuda, T., Gazdar, A.F., 2003. The TP53 gene, tobacco exposure, and lung cancer. *Human Mutation.* **21,** 229-239.

Travis, W.D., Brambilla, E., Müller-Hermelink, H.K. and Harris, C.C., 2004. Chapter 1. Tumours of the Lung. *Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart: World Health Organization Classification of Tumours* IARCPress. 9-124.

Travis, W.D., Colby, T.V., Corrin, B., Shimosato, W. and Brambilla, E., 1999. *Histological typing of lung and pleural tumours.* 3rd ed. Berlin: Spreinger-Vorlag.

Travis, W.D., Brambilla, E., Noguchi, M., Nicholson, A., Geisinger, K., Yatabe, Y., Ishikawa, Y., Wistuba, I., Flieder, D.B., Franklin, W., Gazdar, A., Hasleton, P.S., Henderson, D.W., Kerr, K.M., Nakatani, Y., Petersen, I., Roggli, V., Thunnissen, E., Tsao, M., 2012. Diagnosis of Lung Adenocarcinoma in Resected Specimens: Implications of the 2011 International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society Classification. *Archives of Pathology and Laboratory Medicine.*

Travis, W.D., Brambilla, E., Noguchi, M., Nicholson, A.G., Geisinger, K.R., Yatabe, Y., Beer, D.G., Powell, C.A., Riely, G.J., Van Schil, P.E., Garg, K., Austin, J.H., Asamura, H., Rusch, V.W., Hirsch, F.R., Scagliotti, G., Mitsudomi, T., Huber, R.M., Ishikawa, Y., Jett, J., Sanchez-Cespedes, M., Sculier, J.P., Takahashi, T., Tsuboi, M., Vansteenkiste, J., Wistuba, I., Yang, P.C., Aberle, D., Brambilla, C., Flieder, D., Franklin, W., Gazdar, A., Gould, M., Hasleton, P., Henderson, D., Johnson, B., Johnson, D., Kerr, K., Kuriyama, K., Lee, J.S., Miller, V.A., Petersen, I., Roggli, V., Rosell, R., Saijo, N., Thunnissen, E., Tsao, M., Yankelewitz, D., 2011. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **6,** 244-285.

Travis, W.D., Brambilla, E., Riely, G.J., 2013. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **31,** 992-1001.

Travis, W.D., Travis, L.B., Devesa, S.S., 1995. Lung cancer. *Cancer.* **75,** 191-202.

Tsao, S.C., Weiss, J., Hudson, C., Christophi, C., Cebon, J., Behren, A., Dobrovic, A., 2015. Monitoring response to therapy in melanoma by quantifying circulating tumour DNA with droplet digital PCR for BRAF and NRAS mutations. *Scientific Reports.* **5,** 11198.

Tsubota, Y.T., Kawaguchi, T., Hoso, T., Nishino, E., Travis, W.D., 1992. A combined small cell and spindle cell carcinoma of the lung. Report of a unique case with immunohistochemical and ultrastructural studies. *The American Journal of Surgical Pathology.* **16,** 1108-1115.

Tuveson, D.A., Shaw, A.T., Willis, N.A., Silver, D.P., Jackson, E.L., Chang, S., Mercer, K.L., Grochow, R., Hock, H., Crowley, D., Hingorani, S.R., Zaks, T., King, C., Jacobetz, M.A., Wang, L., Bronson, R.T., Orkin, S.H., DePinho, R.A., Jacks, T., 2004. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell.* **5,** 375-387.

Ulivi, P., Foschi, G., Mengozzi, M., Scarpi, E., Silvestrini, R., Amadori, D., Zoli, W., 2013. Peripheral Blood miR-328 Expression as a Potential Biomarker for the Early Diagnosis of NSCLC. *International Journal of Molecular Sciences.* **14,** 10332-10342.

Umemura, S., Mimaki, S., Makinoshima, H., Tada, S., Ishii, G., Ohmatsu, H., Niho, S., Yoh, K., Matsumoto, S., Takahashi, A., Morise, M., Nakamura, Y., Ochiai, A., Nagai, K., Iwakawa, R., Kohno, T., Yokota, J., Ohe, Y., Esumi, H., Tsuchihara, K., Goto, K., 2014. Therapeutic priority of the PI3K/AKT/mTOR pathway in small cell lung cancers as revealed by a comprehensive genomic analysis. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **9,** 1324-1331.

UyBico, S.J., Wu, C.C., Suh, R.D., Le, N.H., Brown, K., Krishnam, M.S., 2010. Lung cancer staging essentials: the new TNM staging system and potential imaging pitfalls. *Radiographics : A Review Publication of the Radiological Society of North America, Inc.* **30,** 1163-1181.

Vanhecke, E., Valent, A., Tang, X., Vielh, P., Friboulet, L., Tang, T., Goubar, A., Li, Y., Robin, A., Behrens, C., Commo, F., Validire, P., Andre, F., Wistuba, I.I., Soria, J.C., Olaussen, K.A., 2013. 19q13-ERCC1 gene copy number increase in non--small-cell lung cancer. *Clinical Lung Cancer.* **14,** 549-557.

Vasioukhin, V., Anker, P., Maurice, P., Lyautey, J., Lederrey, C., Stroun, M., 1994. Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. *British Journal of Haematology.* **86,** 774-779.

Vinayanuwattikun, C., Winayanuwattikun, P., Chantranuwat, P., Mutirangura, A., Sriuranpong, V., 2013. The impact of non-tumor-derived circulating nucleic acids implicates the prognosis of non-small cell lung cancer. *Journal of Cancer Research and Clinical Oncology.* **139,** 67-76.

Virmani, A.K., Fong, K.M., Kodagoda, D., McIntire, D., Hung, J., Tonk, V., Minna, J.D., Gazdar, A.F., 1998. Allelotyping demonstrates common and

distinct patterns of chromosomal loss in human lung cancer types. *Genes, Chromosomes and Cancer.* **21,** 308-319.

von Pawel, J., 2003. The role of topotecan in treating small cell lung cancer: second-line treatment. *Lung Cancer (Amsterdam, Netherlands).* **41 Suppl 4,** S3-8.

Voortman, J., Lee, J.H., Killian, J.K., Suuriniemi, M., Wang, Y., Lucchi, M., Smith, W.I.,Jr, Meltzer, P., Wang, Y., Giaccone, G., 2010. Array comparative genomic hybridization-based characterization of genetic alterations in pulmonary neuroendocrine tumors. *Proceedings of the National Academy of Sciences of the United States of America.* **107,** 13040-13045.

Vousden, K.H. and Lu, X., 2002. Live or let die: the cell's response to p53. *Nature Reviews.Cancer.* **2,** 594-604.

Wagner, P.L., Kitabayashi, N., Chen, Y.T., Saqi, A., 2009. Combined small cell lung carcinomas: genotypic and immunophenotypic analysis of the separate morphologic components. *American Journal of Clinical Pathology.* **131,** 376-382.

Wang, H., Wu, H., Cai, K., Ju, Q., Wang, W., 2012. Phosphatidylinositol 3-kinase could be a promising target in lung cancer therapy. *Journal of B.U.ON.: Official Journal of the Balkan Union of Oncology.* **17,** 729-734.

Wang, L., Hu, H., Pan, Y., Wang, R., Li, Y., Shen, L., Yu, Y., Li, H., Cai, D., Sun, Y., Chen, H., 2014a. PIK3CA mutations frequently coexist with EGFR/KRAS mutations in non-small cell lung cancer and suggest poor prognosis in EGFR/KRAS wildtype subgroup. *PloS One.* **9,** e88291.

Wang, S., Han, X., Hu, X., Wang, X., Zhao, L., Tang, L., Feng, Y., Wu, D., Sun, Y., Shi, Y., 2014b. Clinical significance of pretreatment plasma biomarkers in advanced non-small cell lung cancer patients. *Clinica Chimica Acta; International Journal of Clinical Chemistry.* **430,** 63-70.

Warde, P. and Payne, D., 1992. Does thoracic irradiation improve survival and local control in limited-stage small-cell carcinoma of the lung? A meta-analysis. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **10,** 890-895.

Watanabe, H., Ma, Q., Peng, S., Adelmant, G., Swain, D., Song, W., Fox, C., Francis, J.M., Pedamallu, C.S., DeLuca, D.S., Brooks, A.N., Wang, S., Que, J., Rustgi, A.K., Wong, K.K., Ligon, K.L., Liu, X.S., Marto, J.A., Meyerson, M., Bass, A.J., 2014. SOX2 and p63 colocalize at genetic loci in squamous cell carcinomas. *The Journal of Clinical Investigation.* **124,** 1636-1645.

Wei, Z., Shah, N., Deng, C., Xiao, X., Zhong, T., Li, X., 2016. Circulating DNA addresses cancer monitoring in non small cell lung cancer patients for detection and capturing the dynamic changes of the disease. *SpringerPlus.* **5,** 531-016-2141-5. eCollection 2016.

Weis, B., Schmidt, J., Lyko, F., Linhart, H.G., 2010. Analysis of conditional gene deletion using probe based Real-Time PCR. *BMC Biotechnology.* **10,** 75-6750-10-75.

Weiss, J., Sos, M.L., Seidel, D., Peifer, M., Zander, T., Heuckmann, J.M., Ullrich, R.T., Menon, R., Maier, S., Soltermann, A., Moch, H., Wagener, P., Fischer, F., Heynck, S., Koker, M., Schottle, J., Leenders, F., Gabler, F., Dabow, I., Querings, S., Heukamp, L.C., Balke-Want, H., Ansen, S., Rauh, D., Baessmann, I., Altmuller, J., Wainer, Z., Conron, M., Wright, G., Russell, P., Solomon, B., Brambilla, E., Brambilla, C., Lorimier, P., Sollberg, S., Brustugun, O.T., Engel-Riedel, W., Ludwig, C., Petersen, I., Sanger, J., Clement, J., Groen, H., Timens, W., Sietsma, H., Thunnissen, E., Smit, E., Heideman, D., Cappuzzo, F., Ligorio, C., Damiani, S., Hallek, M., Beroukhim, R., Pao, W., Klebl, B., Baumann, M., Buettner, R., Ernestus, K., Stoelben, E., Wolf, J., Nurnberg, P., Perner, S., Thomas, R.K., 2010. Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Science Translational Medicine.* **2,** 62ra93.

Whale, A.S., Huggett, J.F., Cowen, S., Speirs, V., Shaw, J., Ellison, S., Foy, C.A., Scott, D.J., 2012. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Research.* **40,** e82.

Wheeler, D.L., Dunn, E.F., Harari, P.M., 2010. Understanding resistance to EGFR inhibitors-impact on future treatment strategies. *Nature Reviews.Clinical Oncology.* **7,** 493-507.

Wilbertz, T., Wagner, P., Petersen, K., Stiedl, A.C., Scheble, V.J., Maier, S., Reischl, M., Mikut, R., Altorki, N.K., Moch, H., Fend, F., Staebler, A., Bass, A.J., Meyerson, M., Rubin, M.A., Soltermann, A., Lengerke, C., Perner, S., 2011. SOX2 gene amplification and protein overexpression are associated with better outcome in squamous cell lung cancer. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.* **24,** 944-953.

Wistuba, I.I., Behrens, C., Milchgrub, S., Bryant, D., Hung, J., Minna, J.D., Gazdar, A.F., 1999. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene.* **18,** 643-650.

Wistuba, I.I., Behrens, C., Virmani, A.K., Mele, G., Milchgrub, S., Girard, L., Fondon, J.W.,3rd, Garner, H.R., McKay, B., Latif, F., Lerman, M.I., Lam, S., Gazdar, A.F., Minna, J.D., 2000a. High resolution chromosome 3p allelotyping

of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. *Cancer Research.* **60,** 1949-1960.

Wistuba, I.I., Berry, J., Behrens, C., Maitra, A., Shivapurkar, N., Milchgrub, S., Mackay, B., Minna, J.D., Gazdar, A.F., 2000b. Molecular changes in the bronchial epithelium of patients with small cell lung cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research.* **6,** 2604-2610.

Wistuba, I.I., Gazdar, A.F., Minna, J.D., 2001. Molecular genetics of small cell lung carcinoma. *Seminars in Oncology.* **28,** 3-13.

Wood, D.E., Eapen, G.A., Ettinger, D.S., Hou, L., Jackman, D., Kazerooni, E., Klippenstein, D., Lackner, R.P., Leard, L., Leung, A.N., Massion, P.P., Meyers, B.F., Munden, R.F., Otterson, G.A., Peairs, K., Pipavath, S., Pratt-Pozo, C., Reddy, C., Reid, M.E., Rotter, A.J., Schabath, M.B., Sequist, L.V., Tong, B.C., Travis, W.D., Unger, M., Yang, S.C., 2012. Lung cancer screening. *Journal of the National Comprehensive Cancer Network : JNCCN.* **10,** 240-265.

Woolfenden, S., Zhu, H., Charest, A., 2009. A Cre/LoxP conditional luciferase reporter transgenic mouse for bioluminescence monitoring of tumorigenesis. *Genesis (New York, N.Y.: 2000).* **47,** 659-666.

Xia, S., Huang, C.C., Le, M., Dittmar, R., Du, M., Yuan, T., Guo, Y., Wang, Y., Wang, X., Tsai, S., Suster, S., Mackinnon, A.C., Wang, L., 2015. Genomic variations in plasma cell free DNA differentiate early stage lung cancers from normal controls. *Lung Cancer (Amsterdam, Netherlands).* **90,** 78-84.

Xiang, D., Zhang, B., Doll, D., Shen, K., Kloecker, G., Freter, C., 2013. **Lung cancer screening: from imaging to biomarker**. *Biomarker Research.* **1,** .

Xie, G.S., Hou, A.R., Li, L.Y., Gao, Y.N., Cheng, S.J., 2004. Quantification of plasma DNA as a screening tool for lung cancer. *Chinese Medical Journal.* **117,** 1485-1488.

Xiong, D., Li, G., Li, K., Xu, Q., Pan, Z., Ding, F., Vedell, P., Liu, P., Cui, P., Hua, X., Jiang, H., Yin, Y., Zhu, Z., Li, X., Zhang, B., Ma, D., Wang, Y., You, M., 2012. Exome sequencing identifies MXRA5 as a novel cancer gene frequently mutated in non-small cell lung carcinoma from Chinese patients. *Carcinogenesis.* **33,** 1797-1805.

Yamamoto, H., Shigematsu, H., Nomura, M., Lockwood, W.W., Sato, M., Okumura, N., Soh, J., Suzuki, M., Wistuba, I.I., Fong, K.M., Lee, H., Toyooka, S., Date, H., Lam, W.L., Minna, J.D., Gazdar, A.F., 2008. PIK3CA mutations and copy number gains in human lung cancers. *Cancer Research.* **68,** 6913-6921.

Yim, J., Zhu, L.C., Chiriboga, L., Watson, H.N., Goldberg, J.D., Moreira, A.L., 2007. Histologic features are important prognostic indicators in early stages lung adenocarcinomas. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.* **20,** 233-241.

Yokomizo, A., Tindall, D.J., Drabkin, H., Gemmill, R., Franklin, W., Yang, P., Sugio, K., Smith, D.I., Liu, W., 1998. PTEN/MMAC1 mutations identified in small cell, but not in non-small cell lung cancers. *Oncogene.* **17,** 475-479.

Yokota, J. and Kohno, T., 2004. Molecular footprints of human lung cancer progression. *Cancer Science.* **95,** 197-204.

Yoon, H.I., Park, K.H., Lee, E.J., Keum, K.C., Lee, C.G., Kim, C.H., Kim, Y.B., 2016. Overexpression of SOX2 Is Associated with Better Overall Survival in Squamous Cell Lung Cancer Patients Treated with Adjuvant Radiotherapy. *Cancer Research and Treatment : Official Journal of Korean Cancer Association.* **48,** 473-482.

Yoon, K.A., Park, S., Lee, S.H., Kim, J.H., Lee, J.S., 2009. Comparison of circulating plasma DNA levels between lung cancer patients and healthy controls. *The Journal of Molecular Diagnostics : JMD.* **11,** 182-185.

Yoshida, A., Kohno, T., Tsuta, K., Wakai, S., Arai, Y., Shimada, Y., Asamura, H., Furuta, K., Shibata, T., Tsuda, H., 2013. ROS1-rearranged lung cancer: a clinicopathologic and molecular study of 15 surgical cases. *The American Journal of Surgical Pathology.* **37,** 554-562.

Yoshida, K., Yatabe, Y., Park, J.Y., Shimizu, J., Horio, Y., Matsuo, K., Kosaka, T., Mitsudomi, T., Hida, T., 2007. Prospective validation for prediction of gefitinib sensitivity by epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **2,** 22-28.

Yoshida, Y., Shibata, T., Kokubu, A., Tsuta, K., Matsuno, Y., Kanai, Y., Asamura, H., Tsuchiya, R., Hirohashi, S., 2005. Mutations of the epidermal growth factor receptor gene in atypical adenomatous hyperplasia and bronchioloalveolar carcinoma of the lung. *Lung Cancer (Amsterdam, Netherlands).* **50,** 1-8.

Yoshizawa, A., Motoi, N., Riely, G.J., Sima, C.S., Gerald, W.L., Kris, M.G., Park, B.J., Rusch, V.W., Travis, W.D., 2011. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc.* **24,** 653-664.

Yu, J.B., Decker, R.H., Detterbeck, F.C., Wilson, L.D., 2010. Surveillance epidemiology and end results evaluation of the role of surgery for stage I small cell lung cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **5,** 215-219.

Yuan, P., Kadara, H., Behrens, C., Tang, X., Woods, D., Solis, L.M., Huang, J., Spinola, M., Dong, W., Yin, G., Fujimoto, J., Kim, E., Xie, Y., Girard, L., Moran, C., Hong, W.K., Minna, J.D., Wistuba, I.I., 2010. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PloS One.* **5,** e9112.

Zhang, L., Yu, H., Badzio, A., Boyle, T.A., Schildhaus, H.U., Lu, X., Dziadziuszko, R., Jassem, J., Varella-Garcia, M., Heasley, L.E., Kowalewski, A.A., Ellison, K., Chen, G., Zhou, C., Hirsch, F.R., 2015. Fibroblast Growth Factor Receptor 1 and Related Ligands in Small-Cell Lung Cancer. *Journal of Thoracic Oncology : Official Publication of the International Association for the Study of Lung Cancer.* **10,** 1083-1090.

Zhong, X.Y., Ladewig, A., Schmid, S., Wight, E., Hahn, S., Holzgreve, W., 2007. Elevated level of cell-free plasma DNA is associated with breast cancer. *Archives of Gynecology and Obstetrics.* **276,** 327-331.

Zhu, C.Q., da Cunha Santos, G., Ding, K., Sakurada, A., Cutz, J.C., Liu, N., Zhang, T., Marrano, P., Whitehead, M., Squire, J.A., Kamel-Reid, S., Seymour, L., Shepherd, F.A., Tsao, M.S., National Cancer Institute of Canada Clinical Trials Group Study BR.21, 2008. Role of KRAS and EGFR as biomarkers of response to erlotinib in National Cancer Institute of Canada Clinical Trials Group Study BR.21. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology.* **26,** 4268-4275.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research.* **31,** 3406-3415.