Scientific Research Publishing

Journal of Cancer Therapy, 2023, 14, 317-344
https://www.scirp.org/journal/jct
ISSN Online: 2151-1942
ISSN Print: 2151-1934

Clinical Analysis of the Colorectal Cohort within the Wales Cancer Biobank: A Study of Outcomes and Genetic Screening

Lisa K. Spary[1*], Katie DeLoyde[1], Helen Roberts[2], Fiona Martin[1], Chi Pooi Lee[1], Rachel Butler[2], Malcolm D. Mason[1], Geraldine A. Thomas[1], Alison Parry-Jones[1], Richard A. Adams[1]

[1]Institute of Cancer and Genetics, School of Medicine, University Hospital of Wales Main Building, Cardiff University, Cardiff, UK
[2]All Wales Medical Genetics Service, Institute of Medical Genetics, University Hospital of Wales, Cardiff, UK
Email: *SparyLK@cardiff.ac.uk

How to cite this paper: Spary, L.K., DeLoyde, K., Roberts, H., Martin, F., Lee, C.P., Butler, R., Mason, M.D., Thomas, G.A., Parry-Jones, A. and Adams, R.A. (2023) Clinical Analysis of the Colorectal Cohort within the Wales Cancer Biobank: A Study of Outcomes and Genetic Screening. *Journal of Cancer Therapy*, **14**, 317-344.
https://doi.org/10.4236/jct.2023.147027

Received: May 25, 2023
Accepted: July 16, 2023
Published: July 19, 2023

Abstract

Over the last 12 years, the Wales Cancer Biobank (WCB) has consented to more than 2000 patients with colorectal cancer (CRC). From these patients, clinical data has been collected and patients have been followed through their cancer journey. Clinical data from these patients have been analyzed to identify any correlation between disease grade and outcome. In a small cohort, consisting of 407 patients, WCB has performed genetic analysis on patient primary tumor samples, identifying and characterizing mutations in the *KRAS*, *NRAS*, *BRAF*, *PIK3CA* and *TP53* genes. The majority of patients with CRC who were consented to WCB were male with a mean age of 69 years and received surgery as the primary treatment for their disease. Pathology and disease-free survival data confirmed worse prognoses associated with more advanced disease. Heterogeneity within the primary tumor was explored in a subgroup of patients. Analysis of the *KRAS* and *TP53* genes confirmed that more than 40% of CRC patients who were tested, harbored a genetic mutation within these genes in their primary tumor. Due to the limited sample size tested, most mutations did not show significant differences in disease-free survival, however, mutation of the *BRAF* gene did show a decrease in the disease specific survival, in keeping with the published data. Analysis of the patients diagnosed with CRC within the Biobank has provided us with valuable information on the status of CRC disease and treatment within the Welsh population. Over 12 years of consenting, we have witnessed significant changes in the information that researchers are interested in when sourcing samples for translational research. The development of new drugs that are tailored to the genetics of a cancer is emerging and at WCB we are focusing our collections on samples and data that meet the needs of this ever-evolving field.

Keywords

Colorectal, Cancer, Biobank, Outcome, Genetics, Screening

1. Introduction

Advances in molecular analysis have led to the development of patient-tailored cancer treatments. In the majority of cancer patients, successful treatment is achieved using conventional methods such as surgery, radiotherapy and/or chemotherapy. However, for those patients who do not respond to conventional treatments, the development of tailored treatments that target particular molecular pathways involved in cancer development and progression is increasingly emerging [1]. The promise of targeted therapy has been to be more specific, with improved global cancer control in the individual, thus improving survival, whilst having less impact in terms of toxicities and quality of life. Interestingly, oncologists have been learning to cope with a range of new toxicities which have not been relevant with more traditional chemotherapy agents, with less risk of immunosuppression but higher chances of skin, eye or cardiac effects. To date, the results of these target-specific treatments have been mixed. As single agents, results in some tumors have been disappointing, suggesting that too little is known about the agents and the molecular pathways they are deemed to impact. A notable example is the targeting of the BRAF axis, in various tumor types. Tumors that harbor a BRAF mutation in colorectal cancer commonly have a worse prognosis in the metastatic setting. Agents such as Encorafenib (a BRAF inhibitor), targeted at the altered BRAF protein arising from the mutated gene, appear to have limited single agent effect in colorectal cancer, whilst single agent BRAF inhibition in BRAF mutant metastatic melanoma often has a dramatic effect. When Encorafenib is combined with Cetuximab (an EGFR inhibitor), it on its own is deemed to be ineffective in BRAF mutant colorectal cancer, then we see a positive effect from the dual blockade such that the combination of Cetuximab and Encorafenib is now licensed and found in guidelines as a standard of care second-line therapy treatment for patients with BRAF mutant metastatic colorectal cancer.

In recent years, advances in molecular screening technologies have provided evidence of the genomic alterations that can occur in cancers. These acquired genetic mutations can result in changes in protein expression of the mutated gene leading to abnormal expression levels that can influence the response of the tumor to various treatments available [1]. Understanding these changes in the genome and proteome has resulted in the development of tailored treatments for multiple cancers including colorectal, lung and breast cancer [2] [3] [4].

Globally, CRC is the third most common cancer in males and the second in females, with an estimated 1.4 million cases and 693,900 deaths occurring in 2012 [5]. Within Wales >2300 patients are diagnosed per year and it is the second biggest cancer killer. Approximately, 90% of CRC cases are sporadic

without a family history or genetic predisposition [6]. Extensive studies analyzing the genetics of CRC have identified mutations in the DNA sequence in both oncogenes and tumor-suppressor genes, predominantly the *KRAS* and *p53* genes [7].

One of the key pathways that have been identified in the progression of CRC is the Mitogen Activated Phospho Kinase (MAPK) pathway [8]. The MAPK signaling pathway controls cell proliferation, differentiation and apoptosis [8]. In normal tissue, the activation of the MAPK pathway is controlled through the interaction of an external growth factor such as epidermal growth factor (EGF), with its receptor (EGFR). Mutations that occur within genes that regulate the interaction can result in the inability of cells to switch this mechanism off resulting in continual activation of the pathway.

In advanced colorectal cancer (CRC), one of the main treatments for patients is anti-EGFR therapy (Cetuximab [9] or Panitumumab [2]). Anti-EGFR drugs bind to the EGFR that is present on tumor cells and limits the growth of these cells by inhibiting the *RAS* signaling pathway. The identification of mutations in the *RAS* (KRAS *or NRAS*) gene has determined that patients harboring a mutated *RAS* will unlikely benefit from anti-EGFR therapy [10] and as a result all patients with advanced CRC are now routinely screened for *RAS* mutations. Also, present in the MAPK pathway is the *BRAF* protein, a single point mutation in the gene accounts for most of the cancer associated aberrations in this gene resulting in a valine to glutamine change at residue 600 (V600E). As a consequence of this mutation, the *BRAF* protein is constitutively activated and in advanced disease has a significant detrimental impact on survival. Whilst specific *BRAF* inhibitors have been developed which are effective in other cancers, in metastatic CRC the use of *BRAF* inhibitors has been ineffective with a response rate of approximately 5% confirming that there is still plenty to learn about *BRAF* mutation in mCRC [11] [12] [13].

Over the last 12 years, the Wales Cancer Biobank (WCB) has consented more than 2000 patients with colorectal cancer [14]. From these patients, clinical data has been collected and follow-up has been regularly performed during this period. We have analyzed the data from these patients to identify any correlation between disease grade and outcome. In a cohort of patients, WCB has performed genetic analysis on patient primary tumor samples. Our aim at the WCB is to identify the proportions of colorectal cancers that are currently banked within the WCB that, contain the common mutations that have been identified as potential drivers in CRC and are the targets for developing anti-cancer drugs. We are also able to link the presence of these mutations with clinical parameters and outcomes, identifying any variance in the Welsh cohort in comparison to internationally published data.

2. Materials and Methods

2.1. Patient Recruitment

The Wales Cancer Biobank (WCB) approaches patients in Wales with known or

suspected cancer to ask them to consent to donate biosamples and data for use in future cancer related research. Within WCB, 2217 colorectal cancer patients were recruited between February 2005 and December 2016. All samples, including a retrospective collection of patients' clinical data, were obtained under informed consent and with ethical approval from the Wales Research Ethics Committee 3.

Each patient was followed up every 12 months with the cutoff date of 31st December 2016. Any patients who were listed as alive at this time, but had less than 12 months of follow-up, were excluded from this analysis.

2.2. Tissue Collection

Colorectal tissue samples from surgical specimens were collected and fixed in formalin prior to embedding in paraffin wax. Formalin fixed paraffin embedded (FFPE) tissue sections (4 μm) were cut from each FFPE block and stained with hematoxylin and eosin (H + E) using the Leica Autostainer XL and Leica CU5030 automated coverslip machine. All samples were then verified for tumor content by a certified histopathologist.

2.3. Macrodissection of FFPE Sections

FFPE tumor samples were sectioned and stained with hematoxylin and eosin to determine the regions with the highest tumor nuclei content. Subsequent unstained 10 μm sections were then macrodissected using the annotated tissue sections for guidance prior to DNA extraction.

2.4. DNA Extraction

DNA was extracted using the QIAGEN EZ1 automated system utilizing the EZ1 DNA Tissue Kit and the EZ1 DNA Paraffin Section Card according to the manufacturers' guidelines. Briefly, tissue was incubated in 180 μl ATL buffer plus 20 μl proteinase K at 56˚C for 1 hour, then 90˚C for 1 hour, before DNA was extracted using the QIAGEN EZ1 BioRobot automated system utilizing the EZ1 DNA Tissue Kit and the EZ1 DNA Paraffin Section Card, according to the manufacturer's guidelines. DNA was eluted in 50 μl of EZ1 elution buffer. The quantification and the purity of the extracted DNAs were measured using the NanoDrop 8000 spectrophotometer. Approximately 20 ng of DNA was required for each PCR reaction.

2.5. PCR Amplification

Initial PCR amplification reactions were performed in 25 ul volumes in Megamix Gold buffer, with 20 ng of DNA template. Primers were designed in-house and were used at 10 pMol and 20 pMol per reaction for downstream pyrosequencing and Sanger sequencing respectively. Each PCR reaction was initially denatured at 95˚C for 10 min, followed by either 36 (*TP53*, *NRAS* and *BRAF*) or 38 (*KRAS* and *PIK3CA*) cycles of 95˚C for 30 sec, 60˚C for 30 sec (59˚C for

PIK3CA) and 72˚C for 30 secs. The final extension step was 10min at 72˚C. Prior to Sanger sequencing, amplification products were checked by gel electrophoresis to confirm amplification and check for contamination of the non-template control.

2.6. Pyrosequencing

Prior to the sequencing reaction, PCR products were cleaned using streptavidin sepharose beads and denatured on the pyromark wash station according to the manufacturer's instructions. Results were analyzed using the Q96 Pyromark software for sequence changes in specific gene regions. Analysis of the genes *KRAS*, *NRAS*, *PIK3CA* and *BRAF* were performed by pyrosequencing using the QIAGEN PyroMark Q96 ID according to the manufacturers' guidelines. Sequencing primers were designed in-house. A wild-type, a mutation-positive and a non-template control were included on each run.

2.7. Sanger Sequencing

Mutation analysis of the *TP53* gene was performed by Sanger sequencing using ABI's (Life Technologies) Big Dye Terminator v1.1 system. PCR products were first cleaned using Agencourt's paramagnetic bead technology (AMPure) then 1μl was carried through to the sequencing reaction. The sequencing reaction proceeded as follows; Initial denaturation at 94˚C for 2 min, followed by 25 cycles of 94˚C for 10 sec, 50˚C for 5 sec and 60˚C for 4 min. Sequencing products were cleaned using Agencourt's paramagnetic bead technology (CleanSeq) and run on the ABI 3730 automated DNA sequencer using POP-7 polymer. Sanger sequence traces were manually analyzed against the reference sequence (NM_000546.4) in Mutation Surveyor (SoftGenetics).

2.8. Outcomes

The primary outcomes for the entire colorectal cancer cohort were disease specific survival and time to relapse. Disease specific survival was defined by either survival or by those that died without evidence of their colorectal disease. A relapse was defined as a recurrence of colorectal disease after undergoing curative resection for colorectal cancer.

For the cohort that underwent genetic screening, the primary outcome was to identify the proportions of colorectal cancers that contain the common mutations that have been identified in CRC and link the presence of these mutations with clinical presentation and outcome.

2.9. Statistics

Statistical analysis was performed on the entire colorectal cancer cohort. Kaplan-Meier analysis was used to describe time to event data (disease specific survival and time to relapse). Follow-up time was measured from the date of surgery to the date of the last follow-up, or death/relapse. The Log-rank test was

used to determine statistical significance between survival curves. The 2 and 5-year survival rates quoted are reported alongside a standard error (SE). The level of significance for all tests was set to $p < 0.05$.

Factors which had a p value < 0.2 during univariate analysis were then entered into multivariate analysis, conducted using Cox Regression, using a forward sequential approach. Hazard ratios (HR) resulting from Cox Regression are reported alongside a 95% CI (confidence interval). If two or more variables demonstrated multicollinearity, then the variable with the lowest p value, or the higher clinical relevance, was included in the model.

3. Results

Between February 2005 and November 2016, 2217 patients with colorectal cancer were recruited to the WCB (~10% of the overall CRC population over 11 years). Of these, 97 patients (4%) were diagnosed post-procedure with a benign tumor and 20 patients with a neuroendocrine tumor. One hundred and thirty-nine (6%) patients had less than 12 months follow-up and were listed as alive (**Figure 1**). These patients were excluded from the analysis (**Figure 1**). In addition, ten patients (<1%) were also excluded as they had been lost to follow-up (**Figure 1**). The resulting 2005 patients with colorectal cancer that met all the eligibility criteria were included in the final analysis (**Figure 1**). Median follow-up was 36 months, ranging from 12 - 225 months. As of 31st December 2016, 1385 (69%) patients were listed as alive. These patients had a median follow-up time of 41 months (range 12 - 225).

3.1. Patient Demographics

Baseline characteristics are shown in **Table 1A**. The majority of patients recruited to the WCB were male (60%). The mean age was 69 years (range 23 - 96, standard deviation [SD] = 11.1). Within the colorectal cohort, 99% of patients had a diagnosis of adenocarcinoma, with the remainder listed as "other cancer" (1%). At presentation, 8% of patients had distant metastasis, compared to 15% of patients with synchronous disease at presentation from Welsh Cancer Registry data 2011 [15].

Patients recruited to WCB showed a varied classification of disease. Based on Dukes stage, where recorded (n = 1065 [53%]), the majority of patients (21%) were classified with Stage II disease (n = 415, **Table 1A**), which is representative of the national percentage 21% - 23% (15). Based on the TNM classification (available for n = 1798 [90%]), 61% of patients (n = 1229) presented with T3 disease (**Table 1A**).

3.2. Treatment

Details of treatments received are presented in **Table 1B**. Within the WCB colorectal cohort, 51% of patients (n = 1027) did not receive any form of chemotherapy (CT) treatment (neoadjuvant CT or adjuvant CT) and 5% (n = 107) of patients were treated with both neoadjuvant CT and adjuvant CT (see supplementary **Table S1**).

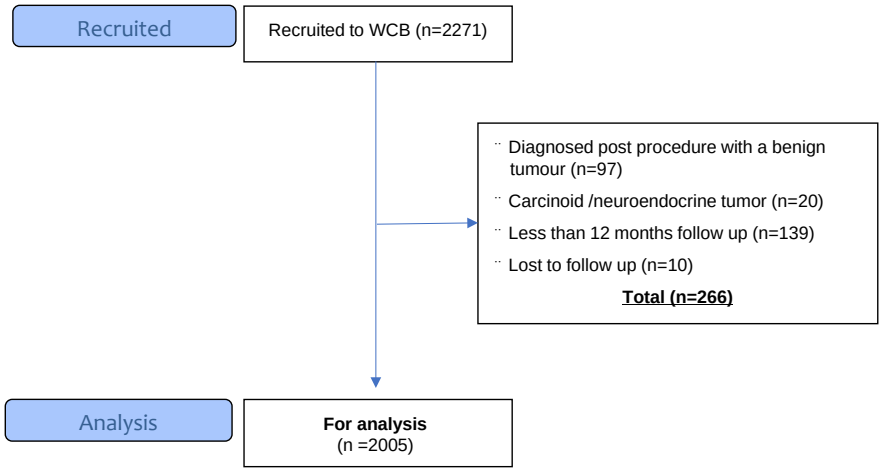

Figure 1. Colorectal recruitment process for the Wales Cancer Bank (WCB). WCB has been routinely consenting patients that have been diagnosed with colorectal cancer (CRC) since 2005. Patients are consented by members of the clinical team and consent allows access to surplus tissue from biopsy and resected tissues, blood samples and clinical data that is recorded in the patients' notes.

Table 1. (A) Baseline characteristics and treatment details; (B) Treatment details.

(A)

	n (%)*		
Demographics	**All patients (n = 2005)**	**Disease-free at death /alive (n = 1499)****	**Disease related death (n = 376)****
Gender			
Female	805 (40)	621 (41)	137 (36)
Male	1200 (60)	878 (59)	239 (64)
Diagnosis			
Adenocarcinoma	1977 (99)	1485 (99)	367 (98)
Other Cancer	28 (1)	14 (<1)	9 (2)
Operation			
Biopsy	108 (5)	36 (2)	50 (13)
Other	34 (2)	14 (<1)	19 (5)
Resection	1863 (93)	1449 (97)	307 (82)
Age			
≤70	1008 (50)	774 (52)	173 (46)
>70	997 (50)	725 (48)	203 (54)
Metastasis at presentation	139 (7)	74 (5)	65 (17)
Staging			
Dukes stage			
A	209 (10)	190 (13)	15 (4)
B	415 (21)	342 (23)	56 (15)
C1	358 (18)	236 (16)	96 (26)
C2	76 (4)	41 (3)	30 (8)
D	7 (<1)	2 (<1)	5 (1)

Continued

T stage			
T1	106 (5)	93 (6)	7 (2)
T2	341 (17)	300 (20)	28 (7)
T3	1229 (61)	919 (61)	242 (64)
T4	122 (6)	78 (5)	26 (7)
N stage			
N0	952 (48)	823 (55)	87 (23)
N1	462 (23)	335 (22)	93 (25)
N2	289 (14)	158 (11)	104 (28)
M stage			
M0	284 (14)	197 (13)	63 (17)
M1	104 (5)	51 (3)	39 (10)
Treatment			
Surgery only	971 (48)	822 (55)	109 (29)
Surgery and CT	700 (35)	474 (32)	172 (46)
Surgery and RT	51 (3)	28 (2)	16 (4)
Surgery and RT and CT	226 (11)	149 (10)	59 (16)
Follow up			
Dead	620 (31)	114 (8)	0
Alive	1385 (69)	1385 (92)	376 (100)
Cause of death			
Not cancer related	114 (6)	114	0
Cancer related	376 (19)	0	376
Alive	1385 (69)	1385	0
Relapse			
1 relapse	477 (24)	183 (12)	243 (65)
3 relapses			
Nodes examined (median [range])	14 (0 - 66)	14 (0 - 56)	12 (0 - 66)
Positive nodes (median [range])	0 (0 - 31)	0 (0 - 27)	2 (0 - 25)

* Where data ≠ 100 data is missing. ** 130 patients were recorded as having an unknown cause of death. Chemotherapy (CT). Radiotherapy (RT).

(B)

	n (%)*		
Treatment details	**All patients (n = 2005)**	**Disease-free at death/alive (n = 1499)****	**Disease related death (n = 376)****
Neoadjuvant CT	198 (10)	136 (9)	52 (14)
Neoadjuvant RT	219 (11)	157 (11)	50 (14)
Adjuvant CT	855 (43)	558 (38)	227 (62)
Adjuvant RT	67 (3)	23 (2)	28 (8)
Treatment for relapse			
Surgery for relapse^	136 (7)	89 (6)	36 (10)
CT for relapse^	247 (12)	77 (5)	133 (35)
RT for relapse^	47 (2)	13 (<1)	30 (8)

* Where data ≠ 100 data is missing. ** 130 patients were recorded as having an unknown cause of death. ^ % of those who had a relapse. Chemotherapy (CT). Radiotherapy (RT).

Radiotherapy (RT) treatment (neoadjuvant RT or adjuvant RT) was given to 16% of patients (n = 328) (see supplementary Table S2). Combination therapy (neoadjuvant CT and neoadjuvant RT) was given to 8% (n = 166) of patients. Whilst 2% (n = 48) of the cohort were treated with a combination of adjuvant CT and adjuvant RT (see supplementary Table S3 and Table S4).

3.3. Disease Specific Survival

The disease specific survival was determined as the percentage of patients within the WCB CRC cohort who have not died from colorectal cancer. All CRC patients consented to WCB between 2005 and 2016 that had at least 12 months of follow-up were analyzed for disease specific survival. Of the 620 patients that were recorded as having died, 114 (18%) were not cancer related and 376 (61%) were colorectal cancer related. For the remaining 130 (21%) patients, the cause of death was either unknown or not recorded. The two and five-year disease specific survival rates were 87.4% (SE 0.8) and 75.3% (SE 1.3) respectively.

As expected, results confirm that a higher T stage correlates with a worse prognosis ($p < 0.001$, Figure 2A). This was also true for the N stage ($p = 0.001$; Figure 2B). Disease specific survival centered on the Dukes system, also confirms a worse prognosis with a higher Dukes score ($p < 0.001$; Figure 2C).

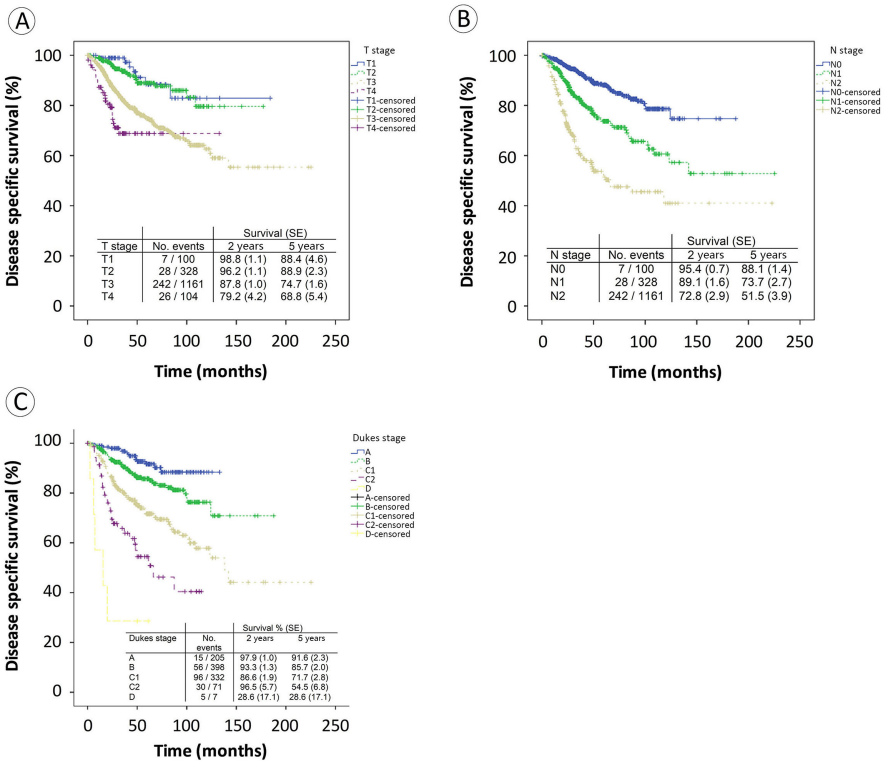

Figure 2. Disease specific survival for colorectal patients recruited to WCB. Disease specific survival times were determined by measuring the time from the date of first treatment to either date of death (caused by their CRC) or the date of last follow-up within 2016. (A) Patients were stratified based on the tumor score of their disease. (B) Patients were stratified on the nodal score (N). Both tumor and nodal score was determined by the TNM classification. (C) Patients were stratified based on Dukes score.

Significant univariate predictors of disease specific survival are listed in Table 2. Multivariate analysis ranked the following variables as independent predictors of an increased risk of cancer related death: >70 years ($p = 0.004$), cancer relapse ($p < 0.001$) and an increased Dukes stage (Table 2).

3.4. Relapse

Within the WCB colorectal cohort, 477 (24%) of patients had a relapse, either a recurrence or secondary metastases (Table 1A). The 2 and 5-year disease free survival were 80.5% (SE 1.0) and 71.8% (SE 1.2) respectively (Figure 3). Univariate predictors of time to relapse are presented in Table 2. After adjustment for treatment type, the Dukes stage continued to have a significant effect on time to relapse (Table 2 and Figure 3). Of those that had a relapse recorded, 63% of patients died compared to 19% of patients that didn't have a relapse recorded ($p < 0.001$).

3.5. Genetic Screening Analysis

Of the 2217 patients with colorectal cancer that were recruited to WCB over the 11 years, tissue samples from 407 patients were selected for genetic testing, analyzing the following genes, *BRAF*, *KRAS*, NRAS, *PI3KCA and TP53*. Patients were chosen based on diagnosis, pathology and length of follow-up available regarding their treatment pathway. Duplicate tumor tissue blocks, representing different tumor regions from the same patient, were included from 11 patients to look for tumor heterogeneity. Any samples that failed the sequencing were excluded from the statistical analysis for that gene. Disease specific survival of patients was analyzed with and without the presence of mutations. Patients that had a cause of death that was not attributed to their CRC disease were excluded from this analysis, along with patients where an unknown cause of death was recorded.

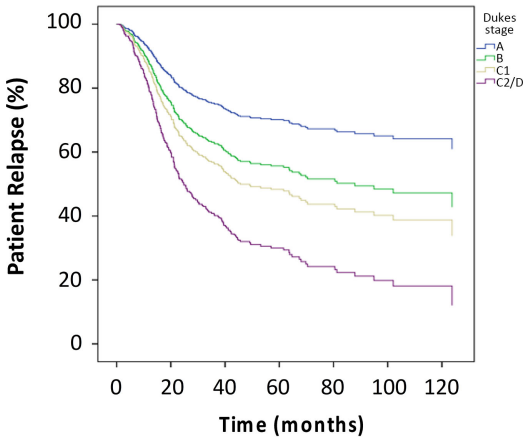

Figure 3. The time taken for Colorectal cohort of patients to relapse. Patient relapses were recorded as either a recurrence or progression of disease or when metastatic lesions were diagnosed after a period of inactive disease. Patients were stratified based on Dukes score.

Table 2. Univariate predictors of disease specific survival and time to relapse.

	Disease specific survival			Time to relapse		
	Survival % n = 2005	Univariate *p* value	Multivariate P value (HR [95% CI])	Survival % n = 2005	Univariate *p* value	Multivariate P value (HR [95% CI])
Gender						
Female	82.0	0.123		76.7	0.269	
Male	78.6			74.4		
Age						
≤70	81.7	0.001	0.004 (1.5 [1.1 - 1.9])	70.6	<0.001	
>70	78.1			80.3		
Diagnosis						
Adenocarcinoma	80.2	0.007		74.7	0.002	
Other cancer	60.9			50.0		
Metastasis at presentation						
No	75.8	<0.001				
Yes	56.7					
Neo adjuvant CT						
No	81.4	0.003		77.2	<0.001	
Yes	72.3			60.9		
Neo adjuvant RT						
No	81.1	0.119		76.9	<0.001	
Yes	75.8			63.5		
Adjuvant CT						
No	86.8	<0.001		83.0	<0.001	
Yes	71.1			66.4		
Adjuvant RT						
No	81.8	<0.001		76.2	0.009	
Yes	45.1			62.5		
Relapse						
No	91.6	<0.001	<0.001 (12.6 [8.9 - 17.8])	-	-	
Yes	43.0					
T stage						
T1	93.0			87.5		
T2	91.5			88.1		
T3	79.2	<0.001		77.4	<0.001	
T4	75.0			58.9		
N stage						
N0	90.4			83.7		
N1	78.3	<0.001		72.9	<0.001	
N2	60.3			58.0		
Dukes stage						
A[R]	92.7			87.9		
B	85.9		0.019 (1.9 [1.1 - 3.5])	78.8		0.047 (1.6 [1.0 - 2.6])
C1	71.1	<0.001	<0.001 (4.2 [2.5 - 7.3])	68.3	<0.001	0.009 (1.9 [1.2 - 3.2])
C2/D	55.1		<0..01 (9.1 [5.0 - 16.8])	56.1		<0.001 (2.9 [1.6 - 5.2])
Treatment						
Surgery only	88.3					
Surgery + CT	73.4					
Surgery + RT	63.6	<0.001				
Surgery + RT + CT	71.6					

[R] Reference category; HR—Hazard ratio; Chemotherapy (CT); Radiotherapy (RT).

3.6. Mutation Analysis

3.6.1. KRAS

Analysis of the *KRAS* gene in the WCB cohort sent for genetic testing (n = 393) identified 163 patients that harbored a mutation (41.2%) (Figure 4Ai). Mutations of the *KRAS* gene were identified in all Dukes scores with an increase in the frequency of *KRAS* mutations being observed as the disease progressed, however, this was not significant (Figure 4Aii). In the WCB cohort, mutations

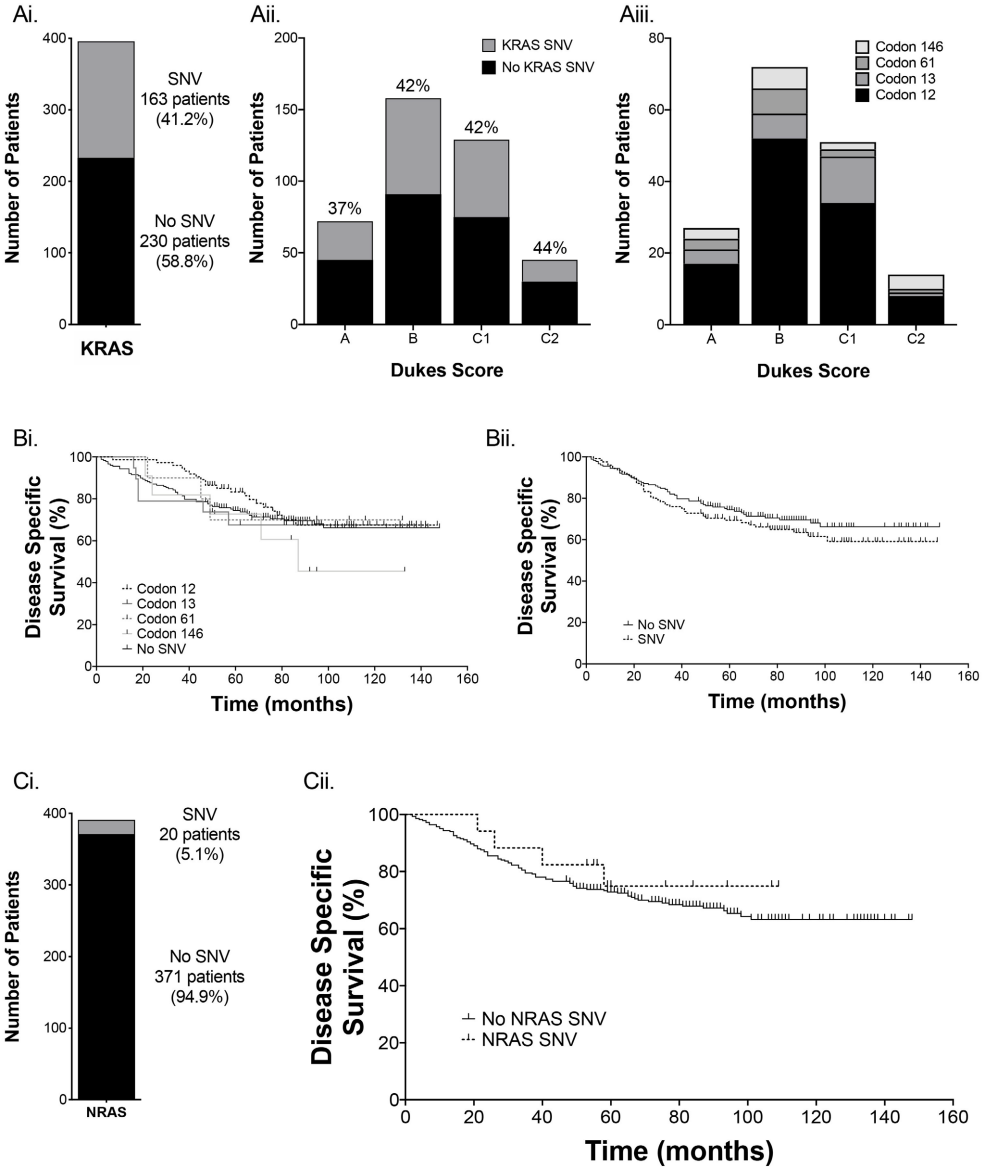

Figure 4. Mutational Analysis of the *RAS* family Members, *KRas* (A and B, n = 393) and *NRas* (C, n = 338). (Ai) Number of patients that harbored a *KRas* single nucleotide variant (SNV). (Aii) Number of patients with a *KRas* SNV stratified by Dukes score. (Aiii) Number of patients with SNV stratified on codon location of mutation versus Dukes score. (Bi) Disease specific survival of patients (n = 301) with a *KRas* SNV versus the codon location. (Bii) Disease specific survival of patients that had a *KRas* SNV versus patients with wildtype (wt) *KRas*. (Ci) Number of patients that harbored a *NRas* SNV. (Cii) Disease specific survival of patients (n = 303) that had a *NRas* mutation versus patients with wildtype (wt) *NRas*.

in codons 12, 13, 61 and 146 were identified, with the majority of the mutations identified occurring within codon 12 (**Figure 4Aiii**). An increase in the number of codon 146 mutations was observed in patients categorized with Dukes C2 disease, whilst higher numbers of codon 13 mutations were observed in the Dukes C1 cohort (**Figure 4Aiii**). Results based on disease specific survival suggest that having a codon 13 mutation is unfavorable, however, this was not significant ($p = 0.824$) (**Figure 4Bi**). Comparisons made between disease specific survival and the presence or absence of a mutation indicated no significant differences ($p = 0.241$) (**Figure 4Bii**). Three out of the 11 patients that were analyzed for tumor heterogeneity contained a *KRAS* mutation there was no evidence of mutational heterogeneity within these cases (data not shown).

3.6.2. NRAS

NRAS mutations were identified in 16 patients within the WCB cohort ($n = 391$; 4.7%) (**Figure 4Ci**). Of these mutations, 50% were located in codon 12, 35% in codon 61 and 15% in codon 13 (data not shown). Interestingly codon 13 mutations were only observed in C1 (only 2 patients) and C2 (only 1 patient) colorectal cancers (data not shown). In the small percentage of the WCB cohort that harbored an *NRAS* mutation, correlation with disease specific survival ($n = 303$) did not identify any disadvantages when compared to those without an *NRAS* mutation (**Figure 4Cii**). No mutations within the *NRAS* gene were identified in the 11 patients that were analyzed for tumor heterogeneity (data not shown).

3.6.3. BRAF

Analysis of the *BRAF* gene in the WCB cohort ($n = 393$) identified 35 patients that harbored the c.1799T > A p.(Val600Glu) mutation, commonly known as V600E (8.9%; **Figure 5Ai**). An increase in the number of *BRAF* mutations in Dukes B, C1 and C2 was observed when compared to Dukes A where no mutations were detected (**Figure 5Aii**). In our cohort, patients harboring a V600E mutation had a significantly worse prognosis ($n = 302$; $p < 0.01$; **Figure 5B**). Notably, 2/14 (14%) Dukes B, 4/11 (36%) Dukes C1, 1/4 (25%) Dukes C2 relapsed. Only one patient that was tested for tumor heterogeneity contained a *BRAF* mutation. Both areas of tumor harbored the same mutation (data not shown).

3.6.4. PIK3CA

Analysis of the *PIK3CA* gene in the WCB cohort ($n = 359$) identified 49 patients that harbored a mutation (13.6%) (**Figure 6Ai**). The majority of *PIK3CA* mutations were observed in Dukes A and B, however there was no significant difference when compared to Dukes C1 and C2 (**Figure 6Aii**). In the WCB population, mutations in both exon 9 and exon 20 were identified (**Figure 6Aiii**). Interestingly exon 20 mutations were not observed in Dukes A colorectal cancers (**Figure 6Aiii**). The disease specific survival ($n = 277$) for patients with an exon 20 mutation suggested an unfavorable prognosis (**Figure 6B**), possibly due to the lack of exon 20 mutations in the Dukes A patients. Results comparing *PIK3CA* mutation versus wildtype (wt) *PIK3CA*, suggests that patients with a mutation

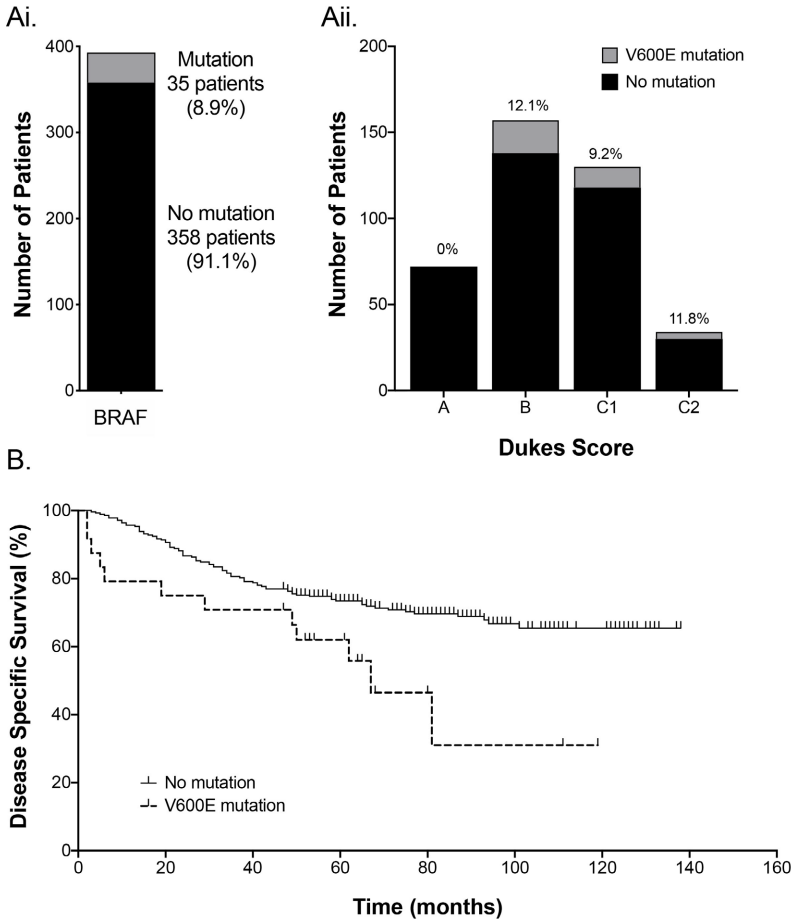

Figure 5. Mutational Analysis of the *BRAF* (V600E) gene (n = 393). (Ai) Number of patients that harbored a *BRAF* V600E mutation. (Aii) Number of patients with a *BRAF* V600E mutation stratified by Dukes score. (B) Disease specific survival of patients (n = 302) with a *BRAF* V600E mutation versus patients with wildtype (wt) *BRAF*.

have an unfavorable prognosis when compared to patients that exhibited no mutation but this was not significant (p = 0.1609; **Figure 6C**). Interestingly, one patient that was tested for tumor heterogeneity harbored a *PIK3CA* mutation in exon 20 in one area of the tumor that was not detected in the other tissue block that was examined (c.3140A > G p.(His1047Arg)).

3.6.5. TP53

A common single nucleotide polymorphism involving the substitution of an arginine for a proline at codon position 72 can be observed in approximately 76.7% of Caucasians. In the WCB cohort (n = 389), 87.40% of patients harbored the common polymorphism c.215C > G p.(Pro72Arg) (data not shown). Many studies have investigated a genetic link between this variation and cancer susceptibility however the results have been inconclusive. Analysis was performed comparing disease specific survival (n = 268) with the presence or absence of the 215C > G SNP and results confirmed that there was no positive or negative effect associated with the SNP (**Figure 7A**). Various *TP53* mutations were observed in

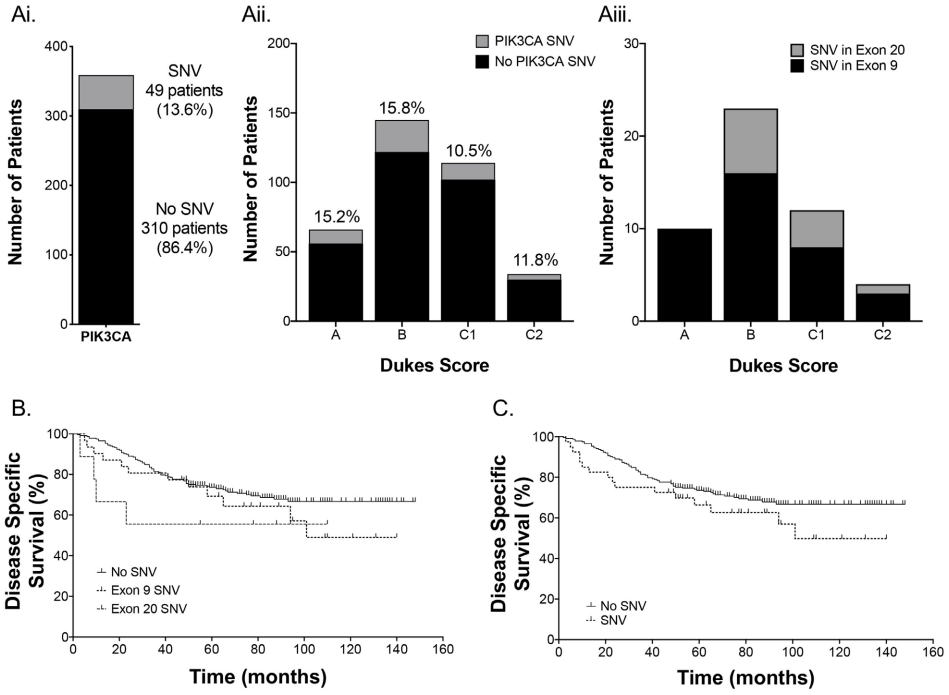

Figure 6. Mutational Analysis of the *PIK3CA* gene (n = 359). (Ai) Number of patients that harbored a *PIK3CA* mutation. (Aii) Number of patients with a *PIK3CA* mutation stratified by Dukes score. (Aiii) Location of the SNV based on Exon versus the Dukes Score of the patient (B) Disease specific survival of patients with either an exon 9 or exon 20 *PIK3CA* mutation versus patients with wildtype (wt) *PIK3CA*. (C) Disease specific survival of patients with a *PIK3CA* mutation versus patients with wildtype (wt) *PIK3CA*.

the WCB cohort (n = 389) with 62% of the patients harboring at least one mutation (Figure 7B). Nine mutations were identified in more than 4 patients, the most prevalent being the c.524G > A p.(Arg175His) mutation found in 16 patients (Figure 7Ci and Table 3). Mutations in these "hotspots" were found to be more prevalent in advanced CRC with 54.5% of patients with Dukes C2 harboring at least one mutation at one of these residues when compared to the whole WCB cohort (Figure 7Cii). Comparisons made between disease specific survival (n = 268) and the presence or absence of a mutation within the WCB cohort suggested patients had a worse prognosis if there was a p53 mutation (p = 0.04) (Figure 7D). Interestingly, five patients out of the 11 tested for tumor heterogeneity harbored a *TP53* mutation in one area of the tumor that was not detected in the other tissue block that was examined (data not shown) suggesting significant multiclonal disease.

4. Discussion

The WCB colorectal cancer patient cohort is representative of the Welsh population, accruing ~10% of patients diagnosed over an 11 years. This powerful dataset including demographic, treatment and outcome data represents a unique resource linked to tissue samples, to support translational research in a disease that has seen little progress in therapies over the last 15 years. Molecular analysis has

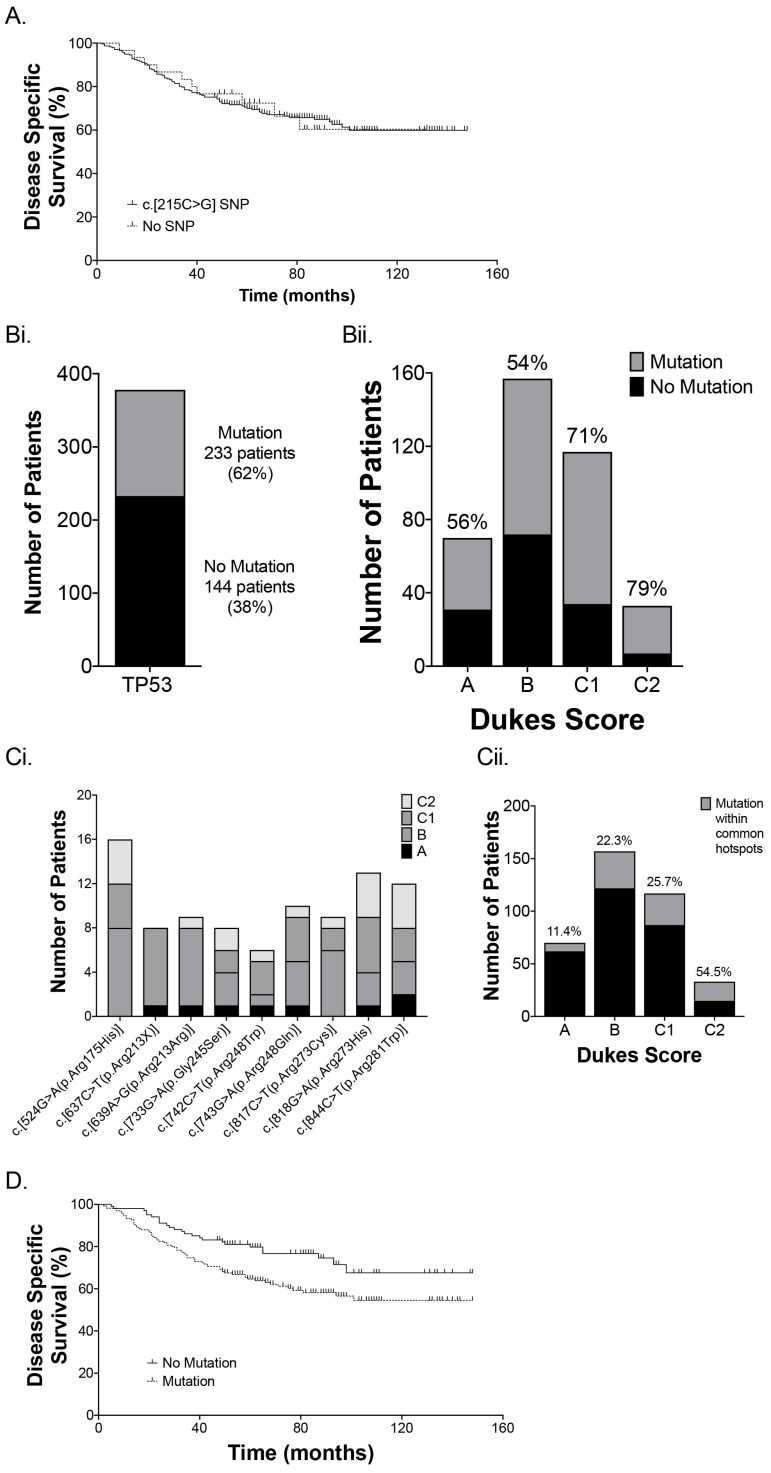

Figure 7. Mutational Analysis of the *TP53* gene (n = 389). (A) Disease specific survival (n = 268) of patients with the common SNP c. [215C > G (p.Pro72Arg)] versus patients without the SNP. (Bi) Number of patients with a *TP53* mutation. (Bii) Number of patients with a *TP53* mutation stratified by Dukes score. (Ci) Number of patients with the common mutation "hotspots" identified with pathogenicity and stratified by Dukes score. (Cii) Percentage of patients within the total cohort that harbor a mutation in the common mutation "hotspots" stratified by Dukes score. (D) Disease specific survival (n = 268) of patients with a *TP53* mutation versus patients with wildtype (wt) *TP53*.

Table 3. List of the mutations commonly detected in CRC Patients within the WCB Cohort.

TP53 Amino Acid Sequence Description	Protein Description	Exon/ Intron	Effect	Transcriptional Activity Class	SNP ID	Clinical Significance
c.[108G > A]	(p.Pro36Pro)	4	silent		rs1800370	benign
c.[375G > A	(p.Thr125Thr)	4	splice		rs55863639	pathogenic
c.[396G > C]	(p.Lys132Asn)	5	missense	non-functional	rs866775781	likely pathogenic
c.[404G > A]	(p.Cys135Tyr)	5	missense	non-functional	rs587781991	likely pathogenic
c.[451C > T]	(p.Pro151Ser)	5	missense	non-functional	rs28943874	likely pathogenic
c.[524G > A]	(p.Arg175His)	5	missense	non-functional	rs28934578	pathogenic
c.[527G > T]	(p.Cys176Phe)	5	missense	partially functional	rs786202962	likely pathogenic
c.[560-1G > A]		5-intron	splice			
c.[584T > C]	(p.Ile195Thr)	6	missense	non-functional	rs587781525	pathogenic
c.[586C > T]	(p.Arg196X)	6	nonsense		rs397516435	pathogenic
c.[635_636delTT]	(p.Phe212fs)	6	frameshift			pathogenic
c.[637C > T]	(p.Arg213X)	6	nonsense		rs397516435	pathogenic
c.[638G > A]	(p.Arg213Gln)	6	missense	non-functional	rs587778720	pathogenic
c.[639A > G]	(p.Arg213Arg)	6	silent		rs1800372	benign
c.[701A > G]	(p.Tyr234Cys)	7	missense	non-functional	rs587780073	likely pathogenic
c.[733G > A]	(p.Gly245Ser)	7	missense	non-functional	rs28934575	pathogenic
c.[734G > A]	(p.Gly245Asp)	7	missense	non-functional	rs121912656	pathogenic
c.[734G > T]	(p.Gly245Val)	7	missense	non-functional	rs121912656	pathogenic
c.[742C > T]	(p.Arg248Trp)	7	missense	non-functional	rs121912651	pathogenic
c.[743G > A]	(p.Arg248Gln)	7	missense	non-functional	rs11540652	pathogenic
c.[811G > T]	(p.Glu271X)	8	missense	partially functional	rs1060501191	uncertain significance
c.[817C > T]	(p.Arg273Cys)	8	missense	non-functional	rs121913343	conflicting interpretations of pathogenicity
c.[818G > A]	(p.Arg273His)	8	missense	non-functional	rs28934576	likely pathogenic
c.[820G > C]	(p.Val274Leu)	8	missense	non-functional	rs1057520005	uncertain significance
c.[844C > T]	(p.Arg281Trp)	8	missense	non-functional	rs28934574	conflicting interpretations of pathogenicity
c.[916C > T	(p.Arg306X)	8	nonsense		rs121913344	pathogenic
c.[919 + 13G > A]		8-intron	splice			
c.[993 + 12T > C]		9-intron	splice		rs1800899	benign

*Mutations listed were found in more than 2 patients within the WCB cohort.

further indicated that on a national level mutation profiling is similar to data provided from clinical trials datasets internationally. At the Wales Cancer Biobank (WCB), the recruitment of patients with colorectal cancer has been underway for more than 13 years. To date, tissue samples from 2217 colorectal patients have been collected along with clinical and follow-up data. Out of these,

samples from 407 patients that had a minimum of 12 months follow-up data were characterized for mutations in the key genes that are known to play a role in cancer development. The analysis of these samples provides insight into the biology of CRC at a national level.

As a cost recovery, not for profit organization, the Wales Cancer Biobank along with other biobanks offers an unrivalled resource to further our understanding of cancer biology, its impact on patient outcomes and our ability to identify targets for future therapeutic intervention. Since the inception of WCB, 19 research groups have applied to WCB for colorectal samples ranging from fresh tumor tissue for the creation of 3D modeling systems for drug discovery, whole blood samples for analyzing the immune cell signatures and plasma and tumor DNA for analyzing temporal changes in circulating biomarkers during treatment.

Statistical analysis of the patient demographics of the WCB colorectal cohort determined that CRC is more prevalent in males within the Welsh population (60% vs 40%) as confirmed using data derived from the Welsh Cancer Intelligence and Surveillance Unit (WCISU) [16]. Compared to other UK countries, the WCB cohort is slightly higher towards men, where the UK average is 55% of newly diagnosed CRC occurring in men [17] [18]. This is also observed when compared to other countries such as the US [19]. Interestingly, within the WCB cohort, only 8% were diagnosed with metastatic colorectal cancer at presentation. This is considerably lower when compared to the UK average, where 23-26% of patients have metastases at diagnosis [20]. This almost certainly relates to the impact of non-removal of the primary tumor in patients who have synchronous metastatic disease, in the WCB cohort of patients, as we have avoided analysis of biopsy only material. As expected, the higher the stage of disease the worse the prognosis. The UK average for net survival for bowel cancer at two- and five-years are 67.9% and 58.7%, respectively [20]. This compares to the disease specific survival rates at two- and five-years for the WCB cohort, of 87.4% and 75.3% respectively. The Wales Cancer Registry WCISU indicates a five-year survival rate of 58.2% [16]. Evidently, factors such as patient consenting are impacted by emergency presentation and synchronous metastatic disease in which no surgical removal of the tumor is planned and this inherently impacts upon prognosis. WCB consented patients' data is heavily reliant on the data recorded in the Cancer Network Information System Cymru (CaNISC) and the Office of National Statistics (ONS). If data was missing or not recorded, patients were excluded from the analysis.

During the lifetime of this research project and over the period these patients have experienced their disease, there have been some significant adaptations in the use of molecular evaluation of tumors to inform clinical practice. Notably, in the adjuvant setting, it has become routine to perform an assessment of microsatellite instability (MSI) in patients with stage II disease. MSI-High tumors are predominantly right sided and often poorly differentiated yet are of better

prognosis and thus may not gain a significant advantage from adjuvant chemotherapy. Nationally MSI testing for all colorectal cancers commenced in June 2019. In the metastatic setting it has become a standard of care to evaluate *RAS* and *BRAF* mutation status and to consider the use of cetuximab or panitumumab in combination with chemotherapy in the first-line setting in those patients with wild type tumors. Further advances have seen the introduction of administering the immunotherapies, Nivolumab plus ipilimumab to patients with MSI-high metastatic CRC. Studies have demonstrated high response rates with increased OS in these patients [21]. The continued screening for genetic aberrations and advances in immunotherapy, especially in metastatic patients, has demonstrated a benefit to patients that in previous years may have had limited treatment options available.

With regard to the mutational analysis, the WCB cohort was significant by its similarities to published data. However, we must accept the limitations of the molecular analysis performed. Analysis of the genes, *KRAS* and *NRAS* were sufficient for the detection of known mutations, but analysis of *BRAF* and *PIK3CA* were limited due to the region covered by the pyrosequencing. Analysis of *TP53* was also limited by the Limit of Detection (LoD) of Sanger sequencing. At the time of analysis, molecular diagnostics was rapidly evolving, although compared to current methods of molecular analysis the techniques utilized was limited, the results are still comparable.

In the WCB cohort, there was a significant decrease in the disease specific survival in patients that harbored a *BRAF* mutation than patients with the wild-type gene. Previous studies have identified *BRAF* mutations in approximately 8% - 15% of all colorectal cancers [22] and have been associated with poor prognosis [12], indicating the importance of *BRAF* mutations in the development and prognosis of CRC. The prominence of a *BRAF* mutation in CRC has previously been shown to affect the response of patients with metastatic disease to targeted therapies such as the anti-EGFR treatments cetuximab or panitumumab [2] [23]. Patients with a wild type *KRAS* gene but harboring a *BRAF* mutation did not respond to the EGFR inhibitors confirming that the *BRAF* mutation plays a critical role in the signaling pathway for EGF [13] [24]. The clinical implications of these findings suggest the need for *BRAF* screening in wild type *KRAS* CRC patients before EGFR inhibitors are administered. In addition, these results demonstrate the further need for research into the pathways involved with the targeted drugs in order to improve the efficacy of these therapies [25].

Mutations in *RAS* family members are frequently found in human cancers including non-small cell lung cancer [26], pancreatic cancer [27] and colorectal cancer [25]. Three *RAS* genes have been identified and although they are functionally distinct they are highly homologous within their genetic sequence [28]. The function of these proteins has a critical role in cell proliferation, survival, and differentiation [28]. The majority of the mutations for the *RAS* genes occur

within codons 12, 13 or 61 and the activating mutations result in constitutive activation leading to a sustained proliferation signal within the cell [28] [29].

Of the three *RAS* family members, studies have shown that mutations within the *KRAS* gene are more frequent in solid tumors, mainly adenocarcinomas, whilst *NRAS* mutations are more prevalent in leukemia [29] [30]. Studies have shown that *KRAS* mutations occur within 40% of CRC whilst *NRAS* mutations occur within 1% - 6% of CRC [31] [32]. Of the 407 CRC cases sent for mutational analysis, 99% of patients had been diagnosed with adenocarcinoma and 41.2% of patients that were tested for *KRAS*, harbored a mutation. In comparison, mutations in *NRAS* were only detected in 4.7% of patients confirming that mutations within the Welsh population are comparable to other populations. Although *NRAS* mutations are rare in CRC, within the WCB CRC cohort the presence of codon 13 mutations was found only in Dukes C1 and C2 but due to low numbers further investigation to determine any correlation will be required. Furthermore, little is known about *NRAS* mutations and their relationship to clinical, pathologic, and molecular features remains uncertain.

As previously mentioned, *KRAS* mutations are more prevalent in codons 12, 13 and 61. In the WCB CRC cohort, the majority of mutations occurred within these regions, however, 4% of the mutations were located with codon 146. In these patients, a single nucleotide change from cytosine to thymine at cDNA position 437 resulted in a protein change from Alanine to Valine. The impact of this mutation is currently unknown and to date, few reports have investigated codon 146 mutations and the clinical relevance of this mutation [32] [33]. Interestingly, when disease specific survival was compared against mutations based on codon location, results suggested that both patients with codon 12 and 13 mutations had a worse prognosis when compared to wt *KRAS*, codon 61 and codon 146 mutations. When disease specific survival was compared for wt *KRAS* and mutated *KRAS*, patients harboring a mutation had an inferior survival but this was not significant. Similar findings have been published by others and therefore the effect of *KRAS* mutations appears to have a greater impact on responses to EGFR treatments than prognosis [25] [34] [35].

Mutations harbored in the *PIK3CA* gene have been implicated in the pathogenesis of multiple cancers including CRC [36]. Somatic mutations in the phosphatidyl 3-kinases (PI3K) family member, *PIK3CA*, result in over activation of the gene which has a role in various cellular processes that can regulate cell proliferation and survival [36]. In CRC, *PIK3CA* mutations are thought to occur within 10% - 30% of cancers with the mutations occurring usually within exon 9 and exon 20 of the gene [37]. In the WCB CRC cohort tested for *PIK3CA* mutation, 13.2% of patients harbored a mutation within exon 9 and/or exon 20. Although patients with a mutated form of *PIK3CA* had an unfavorable prognosis for disease specific survival, this was not significantly different from patients with wt*PIK3CA*. These findings are consistent with other studies that have analyzed *PIK3CA* mutations [38]. A number of studies have reported on the differ-

ent functions of exon 9 and exon 20 of the *PIK3CA* gene and the effects of these in cancer [39] [40] [41]. Analysis of the disease specific survival for *PIK3CA* comparing wt*PIK3CA* with exon 9 and exon 20 mutations separately, suggested that exon 20 mutations resulted in a worse prognosis. This has also been observed in other studies for CRC and other cancers [39] [40] however, due to the low frequency of these mutations the prognostic role of *PIK3CA* mutations, both exon 9 and exon 20, in CRC remains uncertain.

One of the most commonly mutated genes in human cancers is the *p53* gene [42]. It has been reported that more than 50% of CRC will harbor a mutation within the *p53* gene [43]. These mutations are known to drive oncogenic events in CRC, however the mechanisms that *p53* mutations use to exert these events are still unclear. In recent years the development of *p53* targeting agents has been explored but due to the complexity of the oncogenic and biological effects that may occur, few have been translated out of the laboratories [44]. Within the WCB cohort, 62% of patients harbored a *p53* mutation, however most of the mutations detected were present in only one patient confirming the complexity of *p53* mutations. Of the mutations identified, seven mutations were found to be present in more than 3 patients. These mutations are distributed within exon 4-9 which encode the DNA-binding domain [45]. Within this domain, there are six common mutational hotspots, residues R175, G245, R248, R249, R273 and R282 [46]. The most common mutation within the WCB cohort tested, residue R175 (c.524G > A p.(Arg175His)), was identified in 15 patients. Although little is known about the effect this mutation has in CRC, it has been implicated in the activation of c-Met receptor tyrosine kinase in Esophageal squamous cell carcinoma (ESCC) mediating tumor cell invasion [47]. It has also been implicated in endometrial cancers, increasing the invasive phenotypes through activation of the EGFR/PI3K/AKT pathway [48]. All of the mutations observed within the WCB cohort have been identified by various other cancer studies identifying mutations of the *p53* gene. Interestingly *p53* mutations within the common hotspots were identified in patients with more advanced disease (Dukes A 8.8% vs Dukes C2 54.5%) suggesting that *p53* aberrations occur late in tumorigenesis. Studies comparing *p53* mutations in colorectal healthy tissue, adenomas and carcinomas suggest that *p53* mutations develop at late stage adenoma and increase with frequency as carcinomas progress [46]. This has also been documented in other cancers where *p53* mutations have been studied [49] [50]. Studies such as these suggest that *p53* aberrations may be used to determine prognosis, however the clinical significance of *p53* aberrations has long been debated and remains one of the most controversial areas of *p53* research. A review performed in 2010 looking at *p53* mutation and prognosis in multiple cancers suggested that *p53* aberrations within breast, head and neck, liver and haemopoietic cancers were associated with a worse prognosis. This correlation was not conclusive for bladder, brain, lung or ovarian cancer [51].

In addition to the *p53* mutational hotspots, studies have also focused on the

haplotype of *p53* searching for links with cancer prognosis however the results are still inconclusive although most have described a weak association between the SNP c.215C > G p.(Pro72Arg) and an increased risk of CRC [43]. In our cohort, 85.3% of the patients tested had the SNP c.215C > G. When disease-free survival was compared between patients with and without the SNP, we found no significant differences that would suggest that the SNP indicated a worse prognosis.

In the light of precision medicine and improved genetic testing, screening of cancers for gene aberrations has revolutionized the monitoring and treatment of the disease. Our increased knowledge of the mechanisms that tumors use to ensure their progression has enhanced drug development. Genetic alterations in CRC have been studied extensively and with increased sensitivity in screening and detection methods, it continues to advance. Further research on this ever-expanding repertoire of mutations will ensure that future drug development can be tailored to the specifics of the disease rather than a one drug fits all approach.

Acknowledgements

The sample collection by the Wales Cancer Biobank was funded by Welsh Government through Health and Care Research Wales and the research was financed by Cancer Research Wales, CRW Program Grant 2011 (DOI:10.5334/ojb.46). WCB would like to acknowledge all the patients that have donated their samples to the biobank over the years. WCB would also like to acknowledge the Wales NHS, their staff and the continued support that they provide.

Financial Support

The sample collection by the Wales Cancer Biobank was funded by Welsh Government through Health and Care Research Wales. All research performed was financed by Cancer Research Wales, CRW Program Grant 2011.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Verma, M. (2012) Personalized Medicine and Cancer. *Journal of Personalized Medicine*, **2**, 1-14. https://doi.org/10.3390/jpm2010001

[2] Patel, S.B., Gill, D. and Garrido-Laguna, I. (2016) Profile of Panitumumab as First-Line Treatment in Patients with Wild-Type KRAS Metastatic Colorectal Cancer. *OncoTargets and Therapy*, **9**, 75-86. https://doi.org/10.2147/OTT.S68558

[3] Reck, M., Rodríguez-Abreu, D., Robinson, A.G., Hui, R., Csőszi, T., Fülöp, A., *et al.* (2016) Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *The New England Journal of Medicine*, **375**, 1823-1833. https://doi.org/10.1056/NEJMoa1606774

[4] Cho, S.H., Jeon, J. and Kim, S.I. (2012) Personalized Medicine in Breast Cancer: A Systematic Review. *Journal of Breast Cancer*, **15**, 265-272. https://doi.org/10.4048/jbc.2012.15.3.265

[5] Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J. and Jemal, A. (2015) Global Cancer Statistics, 2012. *CA*, **65**, 87-108. https://doi.org/10.3322/caac.21262

[6] Bogaert, J. and Prenen, H. (2014) Molecular Genetics of Colorectal Cancer. *Annals of Gastroenterology*, **27**, 9-14.

[7] Network, C.G.A. (2012) Comprehensive Molecular Characterization of Human Colon and Rectal Cancer. *Nature*, **487**, 330-337. https://doi.org/10.1038/nature11252

[8] Fang, J.Y. and Richardson, B.C. (2005) The MAPK Signalling Pathways and Colorectal Cancer. *The Lancet Oncology*, **6**, 322-327. https://doi.org/10.1016/S1470-2045(05)70168-6

[9] Vale, C.L., Tierney, J.F., Fisher, D., Adams, R.A., Kaplan, R., Maughan, T.S., *et al.* (2012) Does Anti-EGFR Therapy Improve Outcome in Advanced Colorectal Cancer? A Systematic Review and meta-Analysis. *Cancer Treatment Reviews*, **38**, 618-625. https://doi.org/10.1016/j.ctrv.2011.11.002

[10] Spindler, K.L., Pallisgaard, N., Lindebjerg, J., Frifeldt, S.K. and Jakobsen, A. (2011) EGFR Related Mutational Status and Association to Clinical Outcome of Third-Line Cetuximab-Irinotecan in Metastatic Colorectal Cancer. *BMC Cancer*, **11**, Article No. 107. https://doi.org/10.1186/1471-2407-11-107

[11] Barras, D. (2015) *BRAF* Mutation in Colorectal Cancer: An Update. *Biomarkers in Cancer*, **7**, BIC.S25248. https://doi.org/10.4137/BIC.S25248

[12] Kalady, M.F., Dejulius, K.L., Sanchez, J.A., Jarrar, A., Liu, X., Manilich, E., *et al.* (2012) *BRAF* Mutations in Colorectal Cancer Are Associated with Distinct Clinical Characteristics and Worse Prognosis. *Diseases of the Colon & Rectum*, **55**, 128-133. https://doi.org/10.1097/DCR.0b013e31823c08b3

[13] Prahallad, A., Sun, C., Huang, S., Di Nicolantonio, F., Salazar, R., Zecchin, D., *et al.* (2012) Unresponsiveness of Colon Cancer to BRAF(V600E) Inhibition Through Feedback Activation of EGFR. *Nature*, **483**, 100-103. https://doi.org/10.1038/nature10868

[14] Parry-Jones, A. and Spary, L.K. (2018) The Wales Cancer Bank (WCB). *Open Journal of Bioresources*, **5**, 10. https://doi.org/10.5334/ojb.46

[15] Welsh Cancer Intelligence and Surveillance Unit (2022) Cancer Incidence in Wales, 2002-2019. https://phw.nhs.wales/services-and-teams/welsh-cancer-intelligence-and-surveillance-unit-wcisu/cancer-incidence-in-wales-2002-2019/

[16] Welsh Cancer Intelligence and Surveillance Unit (2022) PHW Latest Available Cancer Mortality Official Statistics for Wales for Years 2002 to 2021 by Cancer Type, Sex, Age at Death and Area Disadvantage. WCISU. https://phw.nhs.wales/services-and-teams/welsh-cancer-intelligence-and-surveillance-unit-wcisu/cancer-mortality-in-wales-2002-2021/

[17] Kaur, J. and Poole, J. (2017) Cancer Registration Statistics, England: 2015. Office for National Statistics.

[18] Scotland, I.S.D. (2017) Cancer Incidence in Scotland (2015). NHS.

[19] Siegel, R.L., Miller, K.D., Fedewa, S.A., Ahnen, D.J., Meester, R.G.S., Barzi, A., *et al.* (2017) Colorectal Cancer Statistics, 2017. *A Cancer Journal for Clinicians*, **67**, 177-193.

https://doi.org/10.3322/caac.21395

[20] Cancer Research UK (2018) Bowel cancer Incidence Statistics. Cancer Research UK. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#ref-10

[21] Overman, M.J., Lonardi, S., Wong, K.Y.M., Lenz, H.-J., Gelsomino, F., Aglietta, M., *et al.* (2018) Durable Clinical Benefit with Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *Journal of Clinical Oncology*, **36**, 773-779. https://doi.org/10.1200/JCO.2017.76.9901

[22] Barras, D., Missiaglia, E., Wirapati, P., Sieber, O.M., Jorissen, R.N., Love, C., *et al.* (2017) *BRAF V600E* Mutant Colorectal Cancer Subtypes Based on Gene Expression. *Clinical Cancer Research*, **23**, 104-115. https://doi.org/10.1158/1078-0432.CCR-16-0140

[23] Patel, D.K. (2008) Clinical Use of Anti-Epidermal Growth Factor Receptor Monoclonal Antibodies in Metastatic Colorectal Cancer. *Pharmacotherapy*, **28**, 31S-41S. https://doi.org/10.1592/phco.28.11-supp.31S

[24] Corcoran, R.B., Ebi, H., Turke, A.B., Coffee, E.M., Nishino, M., Cogdill, A.P., *et al.* (2012) EGFR-Mediated Re-Activation of MAPK Signaling Contributes to Insensitivity of BRAF Mutant Colorectal Cancers to RAF Inhibition with Vemurafenib. *Cancer Discovery*, **2**, 227-235. https://doi.org/10.1158/2159-8290.CD-11-0341

[25] Therkildsen, C., Bergmann, T.K., Henrichsen-Schnack, T., Ladelund, S. and Nilbert, M. (2014) The Predictive Value of *KRAS*, *NRAS*, *BRAF*, *PIK3CA* and PTEN for Anti-EGFR Treatment in Metastatic Colorectal Cancer: A Systematic Review and Meta-Analysis. *Acta Oncologica*, **53**, 852-864. https://doi.org/10.3109/0284186X.2014.895036

[26] Chapman, A.M., Sun, K.Y., Ruestow, P., Cowan, D.M. and Madl, A.K. (2016) Lung Cancer Mutation Profile of EGFR, ALK and KRAS: Meta-Analysis and Comparison of never and ever Smokers. *Lung Cancer*, **102**, 122-134. https://doi.org/10.1016/j.lungcan.2016.10.010

[27] Agarwal, A. and Saif, M.W. (2014) KRAS in Pancreatic Cancer. *Journal of the Pancreas*, **15**, 303-305.

[28] Schubbert, S., Shannon, K. and Bollag, G. (2007) Hyperactive Ras in Developmental Disorders and Cancer. *Nature Reviews Cancer*, **7**, 295-308. https://doi.org/10.1038/nrc2109

[29] Fernández-Medarde, A. and Santos, E. (2011) Ras in Cancer and Developmental Diseases. *Genes & Cancer*, **2**, 344-358. https://doi.org/10.1177/1947601911411084

[30] Irahara, N., Baba, Y., Nosho, K., Shima, K., Yan, L., Dias-Santagata, D., *et al.* (2010) *NRAS* Mutations Are Rare in Colorectal Cancer. *Diagnostic Molecular Pathology*, **19**, 157-163. https://doi.org/10.1097/PDM.0b013e3181c93fd1

[31] Amado, R.G., Wolf, M., Peeters, M., Van Cutsem, E., Siena, S., Freeman, D.J., *et al.* (2008) Wild-Type *KRAS* Is Required for Panitumumab Efficacy in Patients with Metastatic Colorectal Cancer. *Journal of Clinical Oncology*, **26**, 1626-1634. https://doi.org/10.1200/JCO.2007.14.7116

[32] Edkins, S., O'Meara, S., Parker, A., Stevens, C., Reis, M., Jones, S., *et al.* (2006) Recurrent KRAS Codon 146 Mutations in Human Colorectal Cancer. *Cancer Biology & Therapy*, **5**, 928-932.

[33] Vaughn, C.P., Zobell, S.D., Furtado, L.V., Baker, C.L. and Samowitz, W.S. (2011) Frequency of *KRAS*, *BRAF* and *NRAS* Mutations in Colorectal Cancer. *Genes, Chromosomes and Cancer*, **50**, 307-312. https://doi.org/10.1002/gcc.20854

[34] Phipps, A.I., Buchanan, D.D., Makar, K.W., Win, A.K., Baron, J.A., Lindor, N.M., *et al.* (2013) *KRAS*-Mutation Status in Relation to Colorectal Cancer Survival: The Joint Impact of Correlated Tumour Markers. *British Journal of Cancer*, **108**, 1757-1764. https://doi.org/10.1038/bjc.2013.118

[35] Phipps, A.I., Limburg, P.J., Baron, J.A., Burnett-Hartman, A.N., Weisenberger, D.J., Laird, P.W., *et al.* (2015) Association between Molecular Subtypes of Colorectal Cancer and Patient Survival. *Gastroenterology*, **148**, 77-87. https://doi.org/10.1053/j.gastro.2014.09.038

[36] Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., *et al.* (2004) High Frequency of Mutations of the PIK3CA Gene in Human Cancers. *Science*, **304**, 554. https://doi.org/10.1126/science.1096502

[37] Karakas, B., Bachman, K.E. and Park, B.H. (2006) Mutation of the *PIK3CA* Oncogene in Human Cancers. *British Journal of Cancer*, **94**, 455-459. https://doi.org/10.1038/sj.bjc.6602970

[38] Mei, Z.B., Duan, C.Y., Li, C.B., Cui, L. and Ogino, S. (2016) Prognostic Role of Tumor PIK3CA Mutation in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Annals of Oncology*, **27**, 1836-1848. https://doi.org/10.1093/annonc/mdw264

[39] De Roock, W., Claes, B., Bernasconi, D., De Schutter, J., Biesmans, B., Fountzilas, G., *et al.* (2010) Effects of KRAS, BRAF, NRAS and PIK3CA Mutations on the Efficacy of Cetuximab Plus Chemotherapy in Chemotherapy-Refractory Metastatic Colorectal Cancer: A Retrospective Consortium Analysis. *Lancet Oncology*, **11**, 753-762.

[40] Lai, Y.L., Mau, B.L., Cheng, W.H., Chen, H.M., Chiu, H.H. and Tzen, C.Y. (2008) *PIK3CA* Exon 20 Mutation Is Independently Associated with a Poor Prognosis in Breast Cancer Patients. *Annals of Surgical Oncology*, **15**, 1064-1069. https://doi.org/10.1245/s10434-007-9751-7

[41] Cathomas, G. (2014) PIK3CA in Colorectal Cancer. *Frontiers in Oncology*, **4**, Article 35. https://doi.org/10.3389/fonc.2014.00035

[42] Olivier, M., Hollstein, M. and Hainaut, P. (2010) TP53 Mutations in Human Cancers: Origins, Consequences and Clinical Use. *Cold Spring Harbor Perspectives in Biology*, **2**, a001008. https://doi.org/10.1101/cshperspect.a001008

[43] Naccarati, A., Polakova, V., Pardini, B., Vodickova, L., Hemminki, K., Kumar, R., *et al.* (2012) Mutations and Polymorphisms in TP53 Gene—An Overview on the Role in Colorectal Cancer. *Mutagenesis*, **27**, 211-218. https://doi.org/10.1093/mutage/ger067

[44] Li, X.L., Zhou, J., Chen, Z.R. and Chng, W.J. (2015) P53 Mutations in Colorectal Cancer—Molecular Pathogenesis and Pharmacological Reactivation. *World Journal of Gastroenterology*, **21**, 84-93. https://doi.org/10.3748/wjg.v21.i1.84

[45] Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) Crystal Structure of a p53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations. *Science*, **265**, 346-355. https://doi.org/10.1126/science.8023157

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

[47] Grugan, K.D., Vega, M.E., Wong, G.S., Diehl, J.A., Bass, A.J., Wong, K.K., *et al.* (2013) A Common p53 Mutation (R175H) Activates c-Met Receptor Tyrosine Kinase to Enhance Tumor Cell Invasion. *Cancer Biology & Therapy*, **14**, 853-859. https://doi.org/10.4161/cbt.25406

[48] Dong, P., Xu, Z., Jia, N., Li, D. and Feng, Y. (2009) Elevated Expression of p53 Gain-of-Function Mutation R175H in Endometrial Cancer Cells Can Increase the Invasive Phenotypes by Activation of the EGFR/PI3K/AKT Pathway. *Molecular Cancer*, **8**, Article No. 103. https://doi.org/10.1186/1476-4598-8-103

[49] Schlomm, T., Iwers, L., Kirstein, P., Jessen, B., Köllermann, J., Minner, S., *et al.* (2008) Clinical Significance of p53 Alterations in Surgically Treated Prostate Cancers. *Modern Pathology*, **21**, 1371-1378. https://doi.org/10.1038/modpathol.2008.104

[50] Olivier, M., Langerød, A., Carrieri, P., Bergh, J., Klaar, S., Eyfjord, J., *et al.* (2006) The Clinical Value of Somatic TP53 Gene Mutations in 1,794 Patients with Breast Cancer. *Clinical Cancer Research*, **12**, 1157-1167. https://doi.org/10.1158/1078-0432.CCR-05-1029

[51] Robles, A.I. and Harris, C.C. (2010) Clinical Outcomes and Correlates of TP53 Mutations and Cancer. *Cold Spring Harbor Perspectives in Biology*, **2**, a001016. https://doi.org/10.1101/cshperspect.a001016

Supplementary Tables

Table S1. Chemotherapy.

Chemotherapy					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00 Neither	1027	51.2	52.6	52.6
	1.00 Both	107	5.3	5.5	58.1
	2.00 Adjuvant only	733	36.6	37.6	95.7
	3.00 Neo adjuvant only	84	4.2	4.3	100.0
	Total	1951	97.3	100.0	
Missing	System	54	2.7		
Total		2005	100.0		

Table S2. Radiotherapy

Radiotherapy					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00 Neither	1677	83.6	86.1	86.1
	1.00 Both	5	0.2	0.3	86.4
	2.00 Adjuvant only	61	3.0	3.1	89.5
	3.00 Neo adjuvant only	204	10.2	10.5	100.0
	Total	1947	97.1	100.0	
Missing	−99.00 Missing	58	2.9		
Total		2005	100.0		

Table S3. Neo-adjuvant Radiotherapy plus Chemotherapy.

NEO_Both_RT_CT					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00 Neither	1713	85.4	91.2	91.2
	1.00 both	166	8.3	8.8	100.0
	Total	1879	93.7	100.0	
Missing	−99.00 Missing	96	4.8		
	System	30	1.5		
	Total	126	6.3		
Total		2005	100.0		

Table S4. Adjuvant Radiotherapy plus Chemotherapy.

Adj_Both_RT_CT					
		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	0.00 Neither	1091	54.4	95.8	95.8
	1.00 both	48	2.4	4.2	100.0
	Total	1139	56.8	100.0	
Missing	−99.00 Missing	78	3.9		
	System	788	39.3		
	Total	866	43.2		
Total		2005	100.0		