Pathology - Research and Practice 215 (2019) 152445

Contents lists available at ScienceDirect

Pathology - Research and Practice

journal homepage: www.elsevier.com/locate/prp

Targeted next generation sequencing reveals a common genetic pathway for colorectal cancers with chromosomal instability and those with microsatellite and chromosome stability

Hersh A. Ham-Karim[a,1], Henry O. Ebili[a,*,2], Kirsty Bradshaw[b], Susan D. Richman[c], Wakkas Fadhil[a], Enric Domingo[d], Ian Tomlinson[e], Mohammad Ilyas[a]

[a] Academic Unit of Pathology and Nottingham Molecular Pathology Node, University of Nottingham, Queen's Medical Centre, UK
[b] Centre for Medical Genetics, Nottingham University Hospitals NHS Trust, City Hospital Campus, UK
[c] Department of Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, Wellcome Trust Brenner Building, St James University Hospital, Leeds, UK
[d] Oxford Centre for Cancer Gene Research and NIHR Comprehensive Biomedical Research Centre, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, UK
[e] Institute of Cancer and Genomic Science, University of Birmingham, Birmingham, UK

ARTICLE INFO

Keywords:
colorectal cancer
genomic instability
tumour ploidy
Next-generation sequencing
Microsatellite and chromosomal stable
Chromosomal instability

ABSTRACT

Introduction: Microsatellite stable sporadic colorectal cancers (CRCs) can be classified as either tumours with chromosomal instability (CIN+) or tumours that are 'Microsatellite and Chromosomal Stable' (MACS). The CIN + tumours are aneuploid whilst MACS are near-diploid; little else is known about their differences. We compared the mutation profiles of CIN + and MACS CRCs.

Method: Targeted Next Generation Sequencing for mutation in 26 driver genes (TruSight-26 kit) was undertaken in 46 CIN + and 35 MACSCRCs. Tumours were compared for mutation frequency, allelic imbalance and clonal heterogeneity.

Results: Mutations were detected in 58% genes and, overall, mutation in driver genes was at expected frequencies. Comparison of classes revealed similar mutation frequencies in most genes and allelic imbalance at*APC* and *TP53*. Differences were seen in mutation frequency in *KRAS* (41% CIN+ *vs 68%* MACS, p = 0.015) and *GNAS* (0% CIN+ *vs* 12% MACS, p = 0.032). Twenty percent CIN + CRCs harboured mutations only in *TP53* - a profile not seen in the MACS tumours ($p = 0.009$). None of the differences were significant after multiple testing corrections.

Conclusions: The mutation profiles of CIN and MACS CRCs are similar. The events allowing aneuploidy (or forcing retention of diploidy) remain unknown.

1. INTRODUCTION

In recent years, a number of different molecular classifications for sporadic colorectal cancers (CRCs) have been produced [1–6]. They vary in terms of the number of groups and the features defining each group although all recognize at least two groups. One group (comprising approximately 10-15% of sporadic CRCs and almost all cancers arising in Lynch Syndrome) are deemed to have Microsatellite Instability (MSI+) and arise due to loss of mismatch repair (MMR) function. MSI + tumours have a characteristic genetic profile with frequent mutations in genes such as *TGFβ1RII, IGFIIR* and *BAX* [7,8]. They are associated with a CpG island methylator (CIMP) phenotype and have a diploid or near-diploid genotype [9].

The group of 85-90% of CRCs which are not MSI + are generally considered to have Chromosomal Instability (CIN+) [10]. The cause of CIN is unknown, but CIN + tumours are characterised by large-scale chromosomal changes such as alteration in whole chromosome number, chromosomal deletion and translocation [11]. CIN + CRCs are characterised by an aneuploid genotype [12] and mutation in genes such as *APC, TP53,* and *FBXW7* [9,13–15].

* Corresponding author at: Academic Unit of Pathology and Nottingham Molecular Pathology Node, University of Nottingham, Queen's Medical Centre, Nottingham, NG7 2UH, UK.
E-mail address: Henry.Ebili@nottingham.ac.uk (H.O. Ebili).
[1] Pharmacy Department, College of Medicine, Komar University of Science and Technology, Chaq-Chaq-Qularaisi, Sulaimani City, Iraq.
[2] Department of Morbid Anatomy and Histopathology, Olabisi Onabanjo University, Ago-Iwoye, Nigeria.

https://doi.org/10.1016/j.prp.2019.152445
Received 29 November 2018; Received in revised form 2 May 2019; Accepted 7 May 2019

However, a group consisting of CRCs which are not MSI + and which have a near-diploid karyotype (and are therefore not CIN+) has been identified. This group has become known as Microsatellite and Chromosome Stable (MACS) tumours although they have previously been reported as 'X-type' [16,17]. Data on the proportion of sporadic CRCs which are MACS-type are varied, with figures ranging between 17% and 45% [15,17,18]. They are reported to display a high rate of CpG island methylator phenotype, similar to that of MSI + tumours [19], they may display a lower rate of loss of APC protein compared to CIN + tumours and they may have a lower rate of loss of expression of MLH1 and BAX proteins compared to MSI + tumours [19].

Data on the clinicopathological features associated with MACS CRCs are conflicting and unclear. Silver et al. [20] demonstrated that MACS CRCs are more likely to be left-sided, Cai et al. [19] showed an association with the right colon whilst Hawkins et al. [21] failed to find any association with tumour site. Tang et al. [22] found that MACS CRC are significantly poorly differentiated and mucinous, and have low TP53 mutation rates. In terms of the prognosis of MACS CRCs, Kakar et al. [18] demonstrated that they had improved survival (compared to CIN +) whilst Hawkins et al [21] reported that MACS CRCs had a poorer prognosis than both MSI and CIN + CRCs.

The recently described Consensus Molecular Classification (CMS) of CRC identifies 4 classes of tumours [5]. The CMS1 class includes the MSI + tumours whilst the CMS4 class is comprised of tumours rich in fibrous stroma. The CMS2 class is characterised by high somatic copy number alteration (SCNA) together with activation of *WNT* and *C-MYC* signalling; this maps closely onto the CIN + group of tumours. The CMS3 class is characterised by low SCNA and frequent *KRAS* mutation; this map closely onto the MACS group of tumours. The CMS classification is based on a number of molecular features which include, but are not restricted to, somatic mutations and therefore a strict genetic classifier does not exist.

Given the stark contrast in karyotype between CIN + and MACS tumours, we hypothesised that there may be one or more genetics events which characterise these groups and which may be responsible for permitting an aneuploid state or forcing retention of a near-diploid state. In order to test this, we compared the mutation profiles of non-MSI + CRCs which had been categorised by flow cytometry as either CIN + or MACS. A targeted next generation sequencing (NGS) approach was used and both sets of tumours were tested for mutation in a panel of 26 cancer-related genes.

Table 1
Clinico-pathological and molecular features of sporadic CIN + and MACS CRC

Features	Categories	MACS-CRCs	CIN-CRCs	*P* value	Adjusted *P* value
Sex	M	19 (54%)	25 (54%)	0.996	1
	F	16 (46%)	21 (54%)		
Age	Median	71	65	-	
	Mean	72	66		
Dukes' stage	A/B	12 (43%)	16 (45%)	0.899	1
	C/D	16 (57%)	20 (55%)		
EMVI	V0	9 (26%)	15 (33%)	0.731	1
	V1	18 (51%)	23 (50%)		
	unknown	8 (23%)	8 (17%)		
Location	Rt side	10 (28%)	12 (26%)	0.803	1
	Lt side	25 (72%)	34 (74%)		
KRAS	Wild	11 (32%)	27 (59%)	0.015	0.150
	Mutant	24 (68%)	19 (41%)		
GNAS	Wild	31 (88%)	46 (100%)	**0.032**	0.213
	Mutant	4 (11%)	0 (0%)		
PIK3CA	Wild	24 (68%)	39 (85%)	0.082	0.410
	Mutant	11 (32%)	7 (15%)		
PTEN	Wild	30 (86%)	44 (96%)	0.230	0.920
	Mutant	5 (14%)	2 (4%)		
FGFR2	Wild	35 (100%)	44 (96%)	0.503	1
	Mutant	0 (0%)	2 (4%)		
CDH1	Wild	34 (98%)	46 (100%)	0.432	1
	Mutant	1 (2%)	0 (0%)		
CTNNB1	Wild	35 (100%)	45 (98%)	1	1
	Mutant	0 (0%)	1 (2%)		
MET	Wild	35 (100%)	45 (98%)	1	1
	Mutant	0 (0%)	1 (2%)		
BRAF	Wild	33 (95%)	41 (89%)	0.693	1
	Mutant	2 (5%)	5 (11%)		
FBXW7	Wild	32 (92%)	40 (87%)	0.725	1
	Mutant	3 (8%)	6 (13%)		
APC	Wild	14 (40%)	21 (46%)	0.611	1
	Mutant	21 (60%)	25 (54%)		
TP53	Wild	10 (29%)	11 (24%)	0.636	1
	Mutant	25 (71%)	35 (76%)		
SMAD4	Wild	29 (83%)	40 (87%)	0.607	1
	Mutant	6 (17%)	6 (13%)		
KIT	Wild	34 (98%)	45 (98%)	1	1
	Mutant	1 (2%)	1 (2%)		
NRAS	Wild	33 (94%)	45 (98%)	0.575	1
	Mutant	2 (6%)	1 (2%)		
***TP53* only**	Negative	37	35	0.009	0.150
	Positive	9	0		

2. MATERIALS AND METHODS

2.1. Clinical Samples

Formalin-fixed paraffin-embedded (FFPE) tissue from sporadic CRCs was retrieved from the archives of the Nottingham University Hospitals Department of Histopathology. All patients had undergone surgery in 2004 or 2005. Cases were selected based on the availability of complete clinico-pathological data and the presence of at least 50% tumour cells in the tumour block. We have previously performed DNA extraction, MSI testing, and ploidy assessment on these CRCs [12,23] and our set of 99 CRCs comprised 7 MSI + tumours, 53 CIN + tumours and 39 MACS tumours. The 7 MSI + CRCs were excluded and another 11 CRCs (7 CIN + and 4 MACS) were excluded based on failing the DNA Quality Control (QC) test during library preparation. Baseline clinico-pathological characteristics of the two groups are listed in Table 1 and we have previously reported that there are no significant differences between them [12,23].

2.2. Next generation sequencing (NGS) library preparation

Mutation profiles were determined using the TruSight tumour kit (Illumina, USA) and samples were run on an Illumina MiSeq DNA analyzer (Illumina, USA). The TruSight Tumour kit offers deep coverage of 26 genes across 175 amplicons (with a minimum of 1000X depth of sequencing and an average depth of 7000X).Each sample underwent a QC step to test for template integrity in accordance with the manufacturer's instructions. PCR based library preparation was carried out in accordance with the manufacturers' instructions. The libraries consist of PCR products of the targeted sequences flanked by common adapters (required for cluster generation and sequencing) and index sequences used to identify individual samples. The libraries for each sample were cleaned up, diluted to a final concentration of 4 nM and then all libraries were pooled into a single tube. Captured libraries were amplified and sequenced as paired-end reads on a MiSeq flow cell. A total of 12 samples were run on each flow cell.

2.3. NGS assay performance

Template derived from FFPE tissue is known to have limitations i.e. the size of the PCR amplicon is limited, the presence of contaminants may inhibit the PCR, there are frequent C→T mutation artefacts due to deamination and there is a higher frequency of spontaneous PCR errors due to degraded template. To assess the utility of the TruSight tumour kit for use in FFPE tissue, we evaluated the frequency of spontaneous mutations and C→T mutation artefacts. Since this platform sequences both strands, we can compare the number of changes occurring in one strand but not seen in the opposite strand and calculate the frequency of

artefact in that target. We also tested the short-term precision (intra-assay variability), the long-term precision (inter-assay variability) and the limit of detection of the assay. For short term precision, one of the samples was tested in 8 replicates in the same run. For long term precision, the same sample was tested on 3 different runs. In each case, both the depth of sequencing and the mutant allele frequencies were evaluated. For limit of detection, a series of dilution experiments were carried out using DNA from two diploid CRC cell lines. DNA from Vaco5 harbouring a BRAF V600E (c.1799 T > A) heterozygous mutation was spiked into DNA extracted from HCT116, which is wild-type at codon 600. DNAs were mixed to produce samples containing mutant alleles at the following percentages; 50%, 12.5%, 6%, 3%, 1.5% and 0.75%.

2.4. QMC-PCR and high resolution melting (HRM) analysis

In order to validate the mutations detected by NGS, the samples were also analysed using the QMC-PCR in conjunction with a high resolution melting (HRM) protocol as previously described [24]. Briefly, the QMC-PCR method is a nested PCR involving a pre-diagnostic multiplex (PDM) reaction and subsequent single specific diagnostic (SSD) reaction. HRM analysis relies on the fact that different DNA sequences display different melting properties (especially when heteroduplexes are formed between wild-type and mutant sequences) which can be distinguished on the basis of fluorescence. The PCR products generated by QMC-PCR were transferred to Light Cycler capillaries (Roche, Germany). The products were loaded into the HR-1 instrument (Idaho Technology, United States) for DNA melting and the HR-1 analysis tool custom software was used to analyze all data. Derivative plots and difference plots were generated after normalizing and temperature shifting the data. To identify mutant-containing samples, the derivative and difference plots were visually inspected and a cut-off threshold of 4% difference in fluorescence (when compared with a known wild-type control) was used as previously described [24,25].

2.5. COLD-PCR

Co-amplification-at-lower-denaturation-temperature-PCR (COLD-PCR) allows the selective enrichment of 'minority alleles' from mixtures of wild-type and mutant sequences [26]. This takes advantage of the fact that heteroduplexes form between wild-type and mutant sequences; these have a lower melting temperature and will therefore undergo denaturation at a lower temperature than homoduplexes. For each DNA sequence, a critical denaturation temperature (Tc) can be identified at which the heteroduplexes (containing equal amounts of both wild-type and mutant sequence) will be completely denatured whilst the homoduplexes (mostly containing the majority wild-type sequence), will still be in double stranded form. The Tc is different for each target sequence and target-specified Tc was determined using a gradient PCR as previously described[25]. COLD-PCR was performed in the SSD step of the QMC-PCR with the following cycling conditions: 95 °C, 5 min; 45 cycles of (95 °C for 15 s; 70 °C for 30 s; target-specified Tc for 20 s; 55.5 °C for 15 s; 72 °C for 15 s), followed by a final extension step of 72 °C for 5 minutes).

2.6. NGS Data analysis

After sequencing, the raw signal data were analysed using MiSeq Reporter v2.1. The pipeline includes signal processing, base calling, quality score assignment, trimming of adapter sequences, filtering for high quality reads, PCR duplicate removal, read alignment to the human genome (hg19) sequence, coverage analysis, and variant calling. After primary data analysis, detected sequence variants (including single nucleotide variants (SNVs) and insertions or deletions (indels)) were assembled in a variant call file (VCF) format generated by the MiSeq Reporter Program. Variant filtering and annotation were performed with Variantstudio™ v2.1analyser. Somatic SNVs and indels were called based on the following criteria: (1) present in both forward and reverse sequencing pools; (2) average read depth of > 500x per pool, (3) > 3% variant frequency in the merged VCF files; (4) minimum Q-score of 20 (phredQ score is a prediction of the possibility of an incorrect base call); (5) not a known germline polymorphism according to dbSNP. For the evaluation of the mutation artefact, criterion (1) and (3) were excluded. For the analysis of the limit of detection, criterion (3) was excluded. The dbSNP reference was used to separate germline from somatic sequence variants.

2.7. Validation of ploidy status by NGS

Both CRC subtypes (i.e. diploid and aneuploid) have allelic imbalance (AI) or loss of heterozygosity. However, CIN (i.e. aneuploid) tumours possess a higher frequency of AI than diploid tumours [22]. Within the limitations of the TruSight 26 tumour kit we sought to validate the ploidy status of MACS and CIN tumour groups by comparing the mean number of altered SNPs (i.e. SNPs with allelic imbalance) and the proportion of informative SNPs which showed AI between CIN and MACS tumour groups.

2.8. Statistical and data analysis

Statistical analyses were performed using a combination of the SPSS software package (version 22), and online chi square and Fisher's exact test (GraphPad, www.graphpad.com/quickcalcs/contingency1/and Social Science Statistics www.socscistatistics.com/tests/) and FDR calculator (www.sdmproject.com/utilities/?show = FDR) software [27–29]. Categorical data were tested for associations using Chi square and a two-sided Fisher's exact tests. The difference in the mean number of altered SNPs was calculated using the independent t test (GraphPad), whilst the proportion of informative SNPs which showed AI between the two groups was calculated using the z-score (VassarStat: a website for statistical computation http://vassarstats.net/propdiff_ind.html) [30]. The Benjamini and Hochberg correction was applied to multiple testing at a false discovery rate of 0.05 (5%) using the online FDR calculator. P values and adjusted P values (for multiple testing) of < 0.05 were considered to be statistically significant.

3. RESULTS

3.1. Performance characteristics of NGS

The overall mean sequencing depth of amplicons was 15509 although amplicon-specific mean sequencing depth varied from 1088X (*TP53* exon 2) to 25401X (*PIK3CA* exon 20). Analysis of the raw data showed an overall mean frequency (of all the samples) of C→T mutation artefacts of 23% and a spontaneous (presumably PCR-induced) error frequency of 8%.

The short-term precision assay showed a mean CV of 12.3% (range 8.6% – 15.3%) for sequencing depth and 2.5% (range 1.6%-4.4%) for MAF. The long-term precision assay showed a mean CV of 10.6% (range 3.2% – 15.1%) for sequencing depth and 2.2 % (range 0.01%-6.1%) for MAF.

A series of spiking experiments was carried out, to determine the limit of detection of mutant allele frequency by the NGS platform. High quality DNA template derived from cell lines was used to obtain optimal data. The limit of detection was 3% of mutant alleles

3.2. Validation of mutations

In total, 247 somatic non-synonymous and likely pathogenic mutations were observed in 15 of the 26 genes in the TruSight tumour panel (Fig. 1 and Supplementary Tables 1 and 2). Of these, 238/247 (96.4%) of the mutations were successfully validated using QMC-PCR with HRM (Fig. 2). The remaining nine mutations were present in the

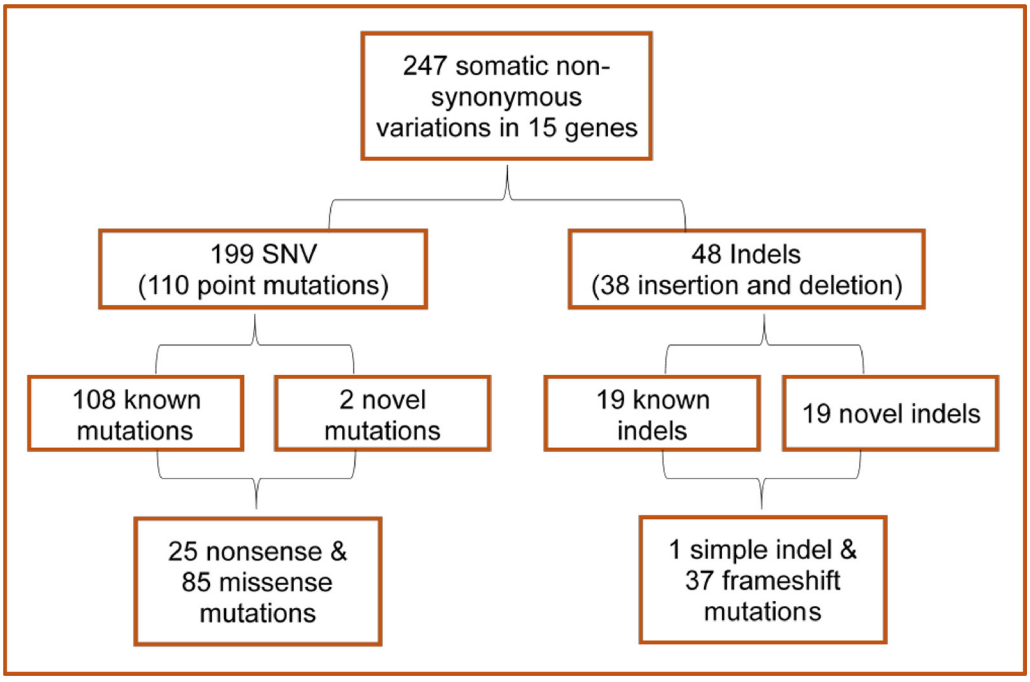

Fig. 1. A schematic showing an analysis of the mutation types found in the 81 CRC samples.

samples at low frequencies and, since the limit of detection of this NGS platform is lower than HRM, we used the COLD-PCR protocol, to enrich for the mutations. This confirmed the presence of the mutations detected by NGS (Fig. 2).

3.3. Allelic imbalance and NGS validation of CIN and MACS tumour group ploidy statuses

Quantification of heterozygous SNPs was used to indicate allelic loss if there is deviation from 50% (outside the range seen in natural assay variation). Based on the maximum CV of 4.4% obtained from the short-term precision assay, and the calculated mean MAF of normal SNPs (49.9%), the normal range for SNPs in the tumour samples was calculated to be 43.3-56.5% for all SNPs.

The numbers of informative SNPs (IS) were 143 for CIN and 108 for MACS. Of these 65/143 (CIN) and 27/108 (MACS) showed AI. There was no significant difference in the mean number of informative SNPs between the CIN and MACS groups (CIN = 3.11 informative SNPs/sample, MACS = 3.09 informative SNPs/sample; mean difference = 0.023, P = 0.95) (Supplementary Tables 3 and 4).

The average numbers of SNPs with AI – and their standard deviations - in the CIN and MACS tumours were calculated in Excel and the resulting values were input into the GraphPad software QuickCalc independent t test calculator. The average numbers of altered SNPs per sample were 1.41 ± 1.31 and 0.77 ± 1.04 in the CIN and MACS groups, respectively. There was a significant difference in the mean number of altered SNPs between CIN and MACS tumour groups (mean difference = 0.64, P value = 0.02).

There was also a significant difference in the proportion of informative SNPs with AI [AI/IS ratios] between the CIN and MACS groups (difference between CIN and MACS proportions = 0.204, z = 3.333, P = 0.0009 (two-tailed)).

3.4. Overall mutation profile

Eighty-one tumour samples were analysed for the presence of

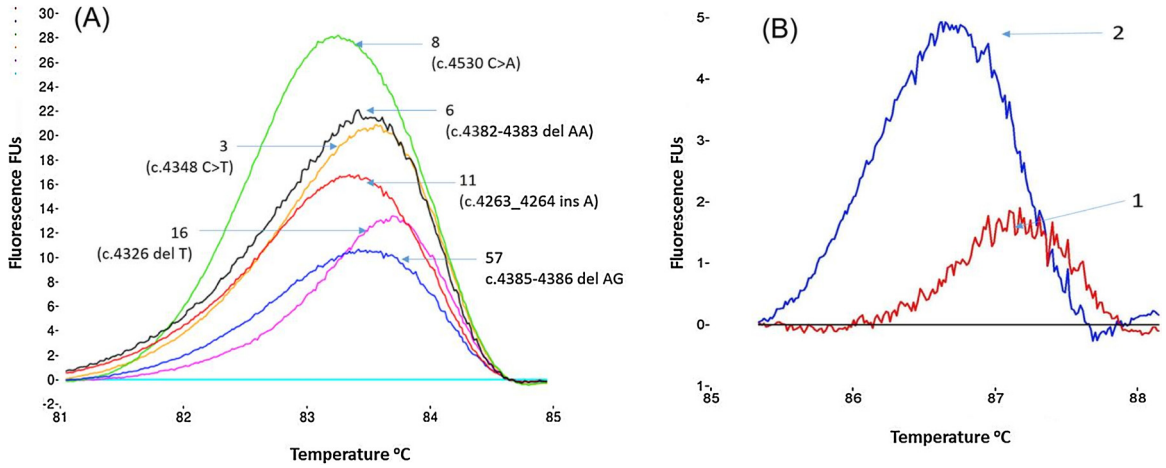

Fig. 2. Validation of the NGS results by HRM analysis. (A) Difference plots obtained from a subset of *APC* mutant samples using QMC-PCR followed by HRM (B) Difference plots obtained from *PTEN* mutant (c.206-209 + 4delGTAA) samples. Validation of this mutation was possible only following minor allele frequency enrichment by COLD PCR.

somatic mutations. Six of 81 (7.4%) tumours showed no mutation in any of the genes included in the panel. The frequencies of mutations detected in the sample set are listed in detail in Table 1, and our data fall within the ranges published in the literature as well as COSMIC and the TGCA (Supplementary Fig. 1 and Supplementary Tables 5 and 6).Co-occurring mutation in each of the set of genes described in the Fearon and Vogelstein model i.e. *APC/(KRAS or BRAF/TP53)* was seen in 27/81 (33.3%) of tumours. The frequency of *APC* mutations (46/81, 57%) was lower than that of *TP53* mutations (60/81, 74%), and 11% (9/81) of CRCs had mutations only in *TP53*. There was a significant positive association between *APC* and *KRAS* mutations (adjusted $P = 0.005$) (Supplementary Table 7). Associations were also seen between *APC* and *PTEN* (adjusted $p = 0.026$), *KRAS* and *PIK3CA* (adjusted $p = 0.002$), *KRAS* and *SMAD4* (adjusted $p = 0.029$) and *PIK3CA* and *GNAS* (adjusted $p = 0.049$). There was no significant association between *APC* and *PIK3CA*. As expected, mutations in *KRAS* and *BRAF* showed a significant negative association (adjusted $p = 0.005$, Fisher's exact test).

3.5. Comparison of mutation profiles between CIN+ and MACS CRCs

NGS analysis revealed similar mutation frequencies within the CIN + and MACS tumour populations (Table 1 and Supplementary Fig. 2). There was no significant difference in the frequency of tumours that were wild-type for all genes in the panel between MACS and CIN + tumours (3/35 MACS *vs*3/46 CIN+, $p = 1.0$). Approximately 20% of CIN+ tumours had a profile of *TP53* only mutation whilst this profile was not seen in the MACS-CRCs (9/46 vs 0/35, $p = 0.009$, adjusted $p = 0.150$). There were also no significant differences in the frequency of *KRAS* and *GNAS* mutation between the two tumour groups following multiple correction testing using the Benjamini and Hochberg correction-based FDR test ($p = 0.015$ and adjusted $p = 0.150$ for *KRAS* and $p = 0.032$ and adjusted $p = 0.213$ for *GNAS*, Table 1).

4. DISCUSSION

4.1. Overall group analysis

In this study we compared the mutation profiles of CIN + and MACS tumours using a targeted NGS approach with a commercially available kit. We found this to be a robust assay producing reproducible data even with low quality template derived from FFPE tissue. The mean sequencing depth overall amplicons was 15509 although there was some variation among amplicons reflecting the range of PCR efficiency for each target. A coefficient of variation (CV) value of 10% is regarded as acceptable for short-term precision and 15% is acceptable for long-term precision. With regards to quantification of the MAF, the assay had mean values well below these limits (CV 2.5% and 2.2% for short term and long-term precision respectively) thereby confirming the reproducibility of the assay.

This NGS platform has a lower limit of detection than QMC-PCR-HRM for mutation detection and thus validation of some of the mutations required mutant enrichment by COLD-PCR-HRM. All 247 somatic mutations detected by NGS were validated showing that there were no false positives. However, spiking experiments with high quality cell-line derived DNA showed a limit of detection for mutant alleles of 3%.This was higher than expected and has implications if this platform is to be used clinically i.e. when considering chemotherapy, resistance mutations may not be detectable if present at a frequency < 3%.

NGS validation of the ploidy statuses of the MACS tumour group compared to the CIN + ve tumours was accomplished using allelic imbalance as previously done [22].

Analysis of the whole dataset (*i.e.* CIN + and MACS combined) reveals frequencies of gene mutation within the published ranges for tumours which are not MSI+ [31–37]. There are, however, some interesting features emerging from our data. Firstly, only one third of the tumours conformed to the Fearon and Vogelstein model (i.e. combined mutation of *APC/(KRAS or BRAF/TP53)* whilst 7% of tumours did not show mutation of any of the genes in this panel. Analysis of COSMIC [38] data and TGCA [39] data showed that the exact mutation frequencies of individual genes that were observed in this study are comparable to those obtained in both series (Supplementary Fig. 1). Thus, most CRCs develop through other pathways. It is uncertain whether these use other gene mutations to deregulate the same signalling pathways (Wnt / MAPK / P53) or whether some CRCs can develop without involvement of these pathways.

The most commonly mutated gene was *TP53 (*74% of cases*)* and 11% of tumours had mutation of *TP53* only (without mutation in any other genes). These data reinforce the importance of *TP53* in CRCs but are interesting inasmuch as there was a lower frequency of APC mutation (57%). Data on *APC* mutation vary with frequency of mutation of this gene reported between 43 – 60% [40,41]. Similar variation is found in the comparison of mutation reported in the COSMIC database (44%) and the TGCA database (80%). Since NGS is known to have a tendency to miss large insertion-deletion (indel) mutations, it is possible that our study has under-called the frequency of *APC* mutation. However, the COSMIC database reports that the 79% of the indels mutations in *APC* tend to be 1 or 2 bases and these would not be expected to be missed. The activation of Wnt signalling is considered an important initiating event in CRC and it is possible that the low frequency of *APC* mutation is may be explained by other events activating Wnt signalling (e.g. mutation of R-spondin, RNF43a).

4.2. Comparison of CIN+ and MACS CRCs

For most of the genes in the panel, there was no difference in frequency of mutation between the two groups. There was a difference in the proportion of tumours showing a profile of *TP53* mutation-only with 9/46 (20%) CIN + tumours having this profile in contrast with 0/35 (0%) MACS tumours. The importance of this is uncertain and, despite an enrichment of tumours with this profile in the CIN + group, there was no difference in the overall frequency of *TP53* mutation between the two groups, contrary to the findings by Tang et al [22]. Similarly, there was no difference in the frequency of *FBXW7* mutation between the two groups. Both genes have previously been associated with the development of aneuploidy [42,43] but this role is not however supported by our study.

There was a slightly greater frequency of *KRAS, PIK3CA and GNAS* mutations in MACS CRCs which was statistically significant for *KRAS* and *GNAS* - although this was lost after correction for multiple testing. It is noteworthy that the CMS 3 class is characterised by common *KRAS* mutation and, since MACS align most closely with CMS 3, this may represent a true association. However, by our data, *KRAS* mutation alone is not a strong discriminator of CIN + mutation from MACS.

Furthermore, the limited number of mutation per sample that could be found using this targeted panel precludes the use of advanced NGS analysis tools such as Somatic Signatures [44], Galaxy tool MutSpec [45], MutationalPatterns [46], and Mutational Signatures in Cancer [47].

In summary, we have demonstrated that this NGS kit and platform make a reliable and sensitive assay for testing of tumours with low quality DNA template. Within the limitations of the TruSight-26 targeted NGS panel our data show that CIN + tumour group has a higher degree of allelic imbalance than the MACS tumours, thereby validating the ploidy status data obtained by flow cytometry. This result is also in line with the established differential characteristics of aneuploidy and diploid tumours [22]. Moreover, our data demonstrated a large degree of overlap in the mutation profiles of CIN + and MACS CRCs. This may indicate a common genetic pathway for these tumour types in the early stages. The events permitting aneuploidy in CIN + CRCs or forcing retention of a near-diploid state in MACS may be unravelled using a more extensive platform such as whole exome or whole genome

sequencing.

Funding

This work was funded by Universities of Nottingham (for MI) and Leeds (for SDR), Wellcome Trust core funding for the WTCHG (090532/Z/09/Z) (for IT) and the Medical Research Council and Cancer Research, UK via the S:CORT consortium for ED. The funding sources played no roles in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Authors' contributions

Experimental design (HAHK, WF, MI), experimentation/data acquisition (HAHK, KM, GC), data analyses and re-analyses (HAHK, KM, SDR, WF, IT, HOE, MI), manuscript drafting (HAHK, MI), manuscript review (HOE, SDR, IT, MI), approval of final manuscript draft (all authors)

Data Statement

All datasets on which the conclusions of this paper rely have been presented in the main manuscript and in the additional supporting files.

Ethics approval and consent to participate

Access to tissues and ethics approval were granted by Nottingham Health Sciences Biobank, which has approval as an IRB from North West—Greater Manchester Central Research Ethics Committee (REC reference: 15/NW/0685).

Consent for publication

No patient consent was needed.

Declaration of conflict of interests

The authors declare no conflict of interest.

Acknowledgements

We would like to thank Gareth Cross for enabling the process of NGS data generation and analyses and Narmeen Ahmed for help with the statistical analyses.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.prp.2019.152445.

References

[1] K. Søreide, B.S. Nedrebø, J.-C. Knapp, T.B. Glomsaker, J.A. Søreide, H. Kørner, Evolving molecular classification by genomic and proteomic biomarkers in colorectal cancer: Potential implications for the surgical oncologist, Surgical Oncology 18 (1) (2009) 31–50.
[2] S. Al-Sohaily, A. Biankin, R. Leong, M. Kohonen-Corish, J. Warusavitarne, Molecular pathways in colorectal cancer, Journal of Gastroenterology and Hepatology 27 (9) (2012) 1423–1431.
[3] J.A. Sanchez, L. Krumroy, S. Plummer, P. Aung, A. Merkulova, M. Skacel, et al., Genetic and epigenetic classifications define clinical phenotypes and determine patient outcomes in colorectal cancer, British Journal of Surgery 96 (10) (2009) 1196–1204.
[4] The Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer, Nature 487 (2012) 330.
[5] J. Guinney, R. Dienstmann, X. Wang, A. de Reyniès, A. Schlicker, C. Soneson, et al., The consensus molecular subtypes of colorectal cancer, Nature Medicine 21 (2015) 1350.
[6] M. Alderdice, S.D. Richman, S. Gollins, J.P. Stewart, C. Hurt, R. Adams, et al., Prospective patient stratification into robust cancer-cell intrinsic subtypes from colorectal cancer biopsies, The Journal of Pathology 245 (1) (2018) 19–28.
[7] A. Duval, M. Reperant, R. Hamelin, Comparative analysis of mutation frequency of coding and non-coding short mononucleotide repeats in mismatch repair deficient colorectal cancers, Oncogene 21 (2002) 8062–8066.
[8] G.A. Calin, R. Gafà, M.G. Tibiletti, V. Herlea, G. Becheanu, L. Cavazzini, et al., Genetic progression in microsatellite instability high (MSI-H) colon cancers correlates with clinico-pathological parameters: A study of the TGRβRII, BAX, hMSH3, hMSH6, IGFIIR and BLM genes, International Journal of Cancer 89 (3) (2002) 230–235.
[9] M. Kouri, A. Laasonen, J.P. Mecklin, H. Jarvinen, K. Franssila, S. Pyrhonen, Diploid predominance in hereditary non-polyposis colorectal carcinoma evaluated by flow cytometry, Cancer 65 (8) (1990) 1825–1829.
[10] W.M. Grady, J.M. Carethers, Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis, Gastroenterology 135 (4) (2008) 1079–1099.
[11] M. Muleris, A. Chalastanis, N. Meyer, M. Lae, B. Dutrillaux, X. Sastre-Garau, et al., Chromosomal Instability in Near-Diploid Colorectal Cancer: A Link between Numbers and Structure, PLoS ONE 3 (2) (2008) e1632.
[12] W. Fadhil, K. Kindle, D. Jackson, A. Zaitoun, N. Lane, A. Robins, et al., DNA content analysis of colorectal cancer defines a distinct 'microsatellite and chromosome stable' group but does not predict response to radiotherapy, International Journal of Experimental Pathology 95 (1) (2014) 16–23.
[13] C. Lengauer, K.W. Kinzler, B. Vogelstein, Genetic instabilities in human cancers, Nature 396 (1998) 643–649.
[14] I.S. Mihaylov, T. Kondo, L. Jones, S. Ryzhikov, J. Tanaka, J. Zheng, et al., Control of DNA Replication and Chromosome Ploidy by Geminin and Cyclin A, Molecular and Cellular Biology 22 (6) (2002) 1868–1880.
[15] M. Pancione, A. Remo, V. Colantuoni, Genetic and Epigenetic Events Generate Multiple Pathways in Colorectal Cancer Progression, Pathology Research International (2012), https://doi.org/10.1155/2012/509348.
[16] M. Linnebacher, C. Ostwald, D. Koczan, T. Salem, B. Schneider, M. Krohn, et al., Single nucleotide polymorphism array analysis of microsatellite-stable, diploid/near-diploid colorectal carcinomas without the CpG island methylator phenotype, Oncology Letters 5 (1) (2013) 173–178.
[17] C. Ostwald, M. Linnebacher, V. Weirich, F. Prall, Chromosomally and microsatellite stable colorectal carcinomas without the CpG island methylator phenotype in a molecular classification, International Journal of Oncology 35 (2) (2009) 321–327.
[18] S. Kakar, G. Deng, V. Sahai, K. Matsuzaki, H. Tanaka, S. Miura, et al., Clinicopathologic characteristics, CpG island methylator phenotype, and BRAF mutations in microsatellite-stable colorectal cancers without chromosomal instability, Arch Pathol Lab Med 132 (6) (2008) 958–964.
[19] G. Cai, Y. Xu, H. Lu, Y. Shi, P. Lian, J. Peng, et al., Clinicopathologic and molecular features of sporadic microsatellite- and chromosomal-stable colorectal cancers, Int J Colorectal Dis 23 (4) (2008) 365–373.
[20] A. Silver, N. Sengupta, D. Propper, P. Wilson, T. Hagemann, A. Patel, et al., A distinct DNA methylation profile associated with microsatellite and chromosomal stable sporadic colorectal cancers, Int J Cancer 130 (5) (2011) 1082–1092.
[21] N.J. Hawkins, I. Tomlinson, A. Meagher, R.L. Ward, Microsatellite-stable diploid carcinoma: a biologically distinct and aggressive subset of sporadic colorectal cancer, Br J Cancer 84 (2001) 232–236.
[22] R. Tang, C.R. Changchien, M.-C. Wu, C.-W. Fan, K.-W. Liu, J.-S. Chen, H.-T. Chien, L.-L. Hsieh, Colorectal cancer without high microsatellite instability and chromosomal instability - an alternative genetic pathway to human colorectal cancer, Carcinogenesis 25 (5) (2004) 841–846, https://doi.org/10.1093/carcin/bgh074.
[23] W. Fadhil, S. Ibrahem, R. Seth, G. Abu-Ali, K. Ragunath, P. Kaye, et al., The utility of diagnostic biopsy specimens for predictive molecular testing in colorectal cancer, Histopathology 61 (6) (2012) 1117–1124.
[24] W. Fadhil, S. Ibrahem, R. Seth, M. Ilyas, Quick-multiplex-consensus (QMC)-PCR followed by high-resolution melting: a simple and robust method for mutation detection in formalin-fixed paraffin-embedded tissue, J Clin Pathol 63 (2) (2010) 134–140.
[25] R. Seth, S. Crook, S. Ibrahem, W. Fadhil, D. Jackson, M. Ilyas, Concomitant mutations and splice variants in KRAS and BRAF demonstrate complex perturbation of the Ras/Raf signalling pathway in advanced colorectal cancer, Gut 58 (2009) 1234–1241.
[26] J. Li, L. Wang, H. Mamon, M.H. Kulke, R. Berbeco, G.M. Makrigiorgos, Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing, Nature Medicine 14 (2008) 579–584.
[27] QuickCalcs GraphPad software, (2018) (Accessed on 20 April 2018), www.graphpad.com/quickcalcs/contingency1/.
[28] Social Science Statistics at http://www.socscistatistics.com/, 2018 (Accessed on 20 April 2018).
[29] FDR online calculator at https://www.sdmproject.com/utilities/?show=FDR, 2018 (Accessed on 20 April 2018)...
[30] VassarStat: a website for statistical computation, (2018) (Accessed on 20 April 2018), http://vassarstats.net/propdiff_ind.html.
[31] R. Fodde, R. Smits, H. Clevers, APC, Signal transduction and genetic instability in colorectal cancer, Nature Reviews Cancer 1 (2001) 55–67.
[32] H. Lamlum, N. Al Tassan, E. Jaeger, I. Frayling, O. Sieber, F.B. Reza, et al., Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q, Human Molecular Genetics 9 (15) (2000) 2215–2221.
[33] M. Lüchtenborg, M.P. Weijenberg, G.M.J.M. Roemen, A.P. de Bruïne, P.A. van den Brandt, M.H.F.M. Lentjes, et al., APC mutations in sporadic colorectal carcinomas from The Netherlands Cohort Study, Carcinogenesis 25 (7) (2004) 1219–1226.

[34] M.P. Macedo, L.D.B. Andrade, R. Coudry, R. Crespo, M. Gomes, B.C.G. Lisboa, et al., Multiple mutations in the Kras gene in colorectal cancer: review of the literature with two case reports, Int J Colorectal Dis. 26 (10) (2011) 1241–1248.
[35] E.R. Fearon, Molecular Genetics of Colorectal Cancer, Annual Review of Pathology: Mechanisms of Disease 6 (1) (2011) 479–507.
[36] W. Qi Li, K. Kawakami, A. Ruszkiewicz, G. Bennett, J. Moore, B. Iacopetta, BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status, Molecular Cancer 5 (1) (2006) 2, https://doi.org/10.1186/1476-4598-5-2DO.
[37] N.I. Fleming, R.N. Jorissen, D. Mouradov, M. Christie, A. Sakthianandeswaren, M. Palmieri, et al., SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer, Cancer Res. 73 (2) (2013) 725–735.
[38] Catalogue of Somatic Mutation in Cancer, (2018) (Accessed on 20 April 2018), https://cancer.sanger.ac.uk/.
[39] The Cancer Genome Atlas, (2018) (Accessed on 20 April 2018), https://cancergenome.nih.gov/.
[40] G. Smith, F.A. Carey, J. Beattie, M.J.V. Wilkie, T.J. Lightfoot, J. Coxhead, et al., Mutations in APC, Kirsten-ras, and p53—alternative genetic pathways to colorectal cancer, PNAS 99 (14) (2002) 9433–9438.
[41] L.J. Gay, P.N. Mitrou, J. Keen, R. Bowman, A. Naguib, J. Cooke, et al., Dietary, lifestyle and clinicopathological factors associated with APC mutations and promoter methylation in colorectal cancers from the EPIC-Norfolk study, The Journal of Pathology 228 (3) (2012) 405–415.
[42] M.L. Bailey, T. Singh, P. Mero, J. Moffat, P. Hieter, Dependence of Human Colorectal Cells Lacking the FBW7 Tumor Suppressor on the Spindle Assembly Checkpoint, Genetics 201 (3) (2015) 885–895.
[43] A.R. Schjolberg, O.P. Clausen, E. Burum-Auensen, P.M. De Angelis, Aneuploidy is associated with TP53 expression but not with BRCA1 or TERT expression in sporadic colorectal cancer, Anticancer Research 29 (1) (2009) 4381–4387.
[44] J.S. Gehring, B. Fischer, M. Lawrence, W. Huber, Somatic Signatures: inferring mutational signatures from single-nucleotide variants, Bioinformatics 31 (2015) 3673–3675.
[45] M. Ardin, V. Cahais, X. Castells, L. Bouaoun, G. Byrnes, Z. Herceg, et al., MutSpec: a galaxy toolbox for streamlined analyses of somatic mutation spectra in human and mouse cancer genomes, BMC Bioinformatics 17 (2016) 170.
[46] F. Blokzijl, J. de Ligt, M. Jager, V. Sasselli, S. Roerink, N. Sasaki, et al., Tissue-specific mutation accumulation in human adult stem cells during life, Nature 538 (2016) 260–264.
[47] M. Díaz-Gay, M. Vila-Casadesús, S. Franch-Expósito, E. Hernández-Illán, J. José Lozano, S. Castellví-Bel, Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples, BMC Bioinformatics 19 (2018) 224, https://doi.org/10.1186/s12859-018-2234-y.