Molecular Diagnosis & Therapy (2019) 23:383–393
https://doi.org/10.1007/s40291-019-00388-z

ORIGINAL RESEARCH ARTICLE

Targeted Next-Generation Sequencing Validates the Use of Diagnostic Biopsies as a Suitable Alternative to Resection Material for Mutation Screening in Colorectal Cancer

Hersh A. Ham-Karim[1,2] · Henry Okuchukwu Ebili[1,3] · Kirsty Manger[4] · Wakkas Fadhil[1] · Narmeen S. Ahmad[5,6] · Susan D. Richman[7] · Mohammad Ilyas[1]

Published online: 11 February 2019

Abstract

Background Mutation testing in the context of neoadjuvant therapy must be performed on biopsy samples. Given the issue of tumour heterogeneity, this raises the question of whether the biopsies are representative of the whole tumour. Here we have compared the mutation profiles of colorectal biopsies with their matched resection specimens.

Methods We performed next-generation sequencing (NGS) analysis on 25 paired formalin-fixed, paraffin-embedded colorectal cancer biopsy and primary resection samples. DNA was extracted and analysed using the TruSight tumour kit, allowing the interrogation of 26 cancer driver genes. Samples were run on an Illumina MiSeq. Mutations were validated using quick-multiplex-consensus (QMC)-polymerase chain reaction (PCR) in conjunction with high resolution melting (HRM). The paired biopsy and resection tumour samples were assessed for presence or absence of mutations, mutant allele frequency ratios, and allelic imbalance status.

Results A total of 81 mutations were detected, in ten of the 26 genes in the TruSight kit. Two of the 25 paired cases were wild-type across all genes. The mutational profiles, allelic imbalance status, and mutant allele frequency ratios of the paired biopsy and resection samples were highly concordant (88.75–98.85%), with all but three (3.7%) of the mutations identified in the resection specimens also being present in the biopsy specimens. All 81 mutations were confirmed by QMC-PCR and HRM analysis, although four low-level mutations required a co-amplification at lower denaturation temperature (COLD)-PCR protocol to enrich for the mutant alleles.

Conclusions Diagnostic biopsies are adequate and reliable materials for molecular testing by NGS. The use of biopsies for molecular screening will enhance targeted neoadjuvant therapy.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40291-019-00388-z) contains supplementary material, which is available to authorized users.

✉ Henry Okuchukwu Ebili
Henry.Ebili@nottingham.ac.uk

[1] Division of Pathology, School of Medicine, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

[2] Department of Medical Laboratory Sciences, College of Health Sciences, Komar University of Science and Technology, Chaq-Chaq-Qualaraisi, Sulaimani, Iraq

[3] Department of Morbid Anatomy and Histopathology, Olabisi Onabanjo University, Ago-Iwoye, Nigeria

[4] Centre for Medical Genetics, Nottingham University Hospitals NHS Trust, City Hospital Campus, Nottingham, UK

[5] Clinical Oncology, University of Nottingham, City Hospital Campus, Nottingham, UK

[6] Kurdistan Institution for Strategic Studies and Scientific Research, Qirga, Sulaimani, KRG, Iraq

[7] Department of Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, Wellcome Trust Brenner Building, St James University Hospital, Leeds, UK

Key Points

The findings from this study have lent credence to the growing notion that diagnostic biopsies are very similar to resection samples at the molecular level.

As such, diagnostic biopsies can be used for molecular testing in place of resection samples.

This creates an opportunity for neoadjuvant therapy and enhances personalised medicine.

1 Introduction

Colorectal cancer (CRC) is the third most common malignancy and a fifth leading cause of cancer deaths worldwide [1]. In the UK, CRC is the fourth most common cancer and the fifth most common cause of cancer deaths, accounting for 10% of all cancer deaths [2]. Recent advances in genome sequencing technologies have enabled greater understanding of the molecular mechanisms of tumourigenesis and aided the identification of clinically relevant biomarkers for diagnosis and personalised therapeutics [3, 4]. Predictive biomarkers are currently used to guide the selection of targeted therapies for personalised medicine. One example of a 'stratified medicine' approach in CRC is tumour assessment for the presence of mutations in the *KRAS* or *NRAS* genes, which predicts a lack of response to epidermal growth factor receptor (EGFR)-targeted antibodies such as panitumumab or cetuximab [4, 5]. Constitutive activation of either *KRAS* or *NRAS* results in excess signalling through the RAS/mitogen-activated protein kinase pathway which cannot be negated by the anti-EGFR monoclonal antibody therapies.

Currently, tumour materials from both biopsy (Bx) and resection specimens (Rx) are recommended for use in the predictive testing of adjuvant targeted therapy response in stage II–III CRC, in the absence of metastatic or recurrent tumour [6]. However, the use of neoadjuvant therapy in patients with CRC is likely to increase, and at present, many predictive biomarkers for neoadjuvant therapy prediction are under study [7, 8]. Whilst neoadjuvant therapy is available for patients with rectal tumours, a clinical trial of neoadjuvant chemotherapy for locally advanced colonic cancer was recently started in the UK and elsewhere [7–10]. In the setting of neoadjuvant therapy, Bx specimens may be the only available specimens to test *KRAS, NRAS* and *BRAF* mutations as recommended for the current standard of care of metastatic CRCs. If the studies on the use of neoadjuvant therapy show desirable outcomes, then the diagnostic Bx specimens may become the only material available for predictive testing in the neoadjuvant settings [11]. CRC develops as a consequence of waves of clonal expansion, resulting from mutations called 'driver mutations' giving a selective advantage [12]. These driver mutations, which are responsible for early clonal sweeps through the adenoma–carcinoma sequence, should therefore be predominantly present in most of the tumour cells and consequently should be present in any Bx samples of an individual tumour.

To confirm whether this is indeed the case and whether diagnostic Bx specimens are appropriate for predictive testing, we have carried out mutation screening of 25 paired diagnostic Bx and their matched Rx. A sensitive next-generation sequencing (NGS) approach was used to assess the presence of mutations in a panel of 26 genes involved with solid tumours.

2 Materials and Methods

2.1 Clinical Samples

Formalin-fixed, paraffin-embedded (FFPE) sporadic CRC tumour blocks were retrieved from the archives of the Nottingham University Hospitals Department of Histopathology. All patients had undergone surgery between 2004 and 2005. Cases were selected based on the availability of clinicopathological data and the presence of at least 50% tumour cells in both Bx and Rx. DNA was extracted using the Qiagen mini kit from 25 cases of paired Bx and Rx samples as previously described [11]. Baseline characteristics are reported in Online Resource Table 1 (see the electronic supplementary material).

2.2 NGS Library Preparation

Mutation profiles were determined using the TruSight tumour kit (Illumina, USA) and samples run on an Illumina MiSeq (Illumina, USA). The TruSight tumour kit offers deep coverage of 26 genes across 175 amplicons (a minimum 1000× coverage, an average of 7000× coverage). Each sample underwent a quality control (QC) step to test for template integrity according to the kit manufacturer's instructions. A polymerase chain reaction (PCR)-based library preparation was carried out in accordance with the manufacturers' instructions. The libraries were cleaned up then diluted to a final concentration of 4 nM before pooling. Captured libraries were amplified and sequenced as paired-end reads on a MiSeq flow cell, with a total of 12 samples being run on each cell.

2.3 NGS Data Analysis

Base calling, quality score assignment and trimming of low-quality reads (using a minimum Q score of 20) were

performed on the MiSeq reporter v2.1 suite. The generated FASTQ files were aligned to the reference genome (hg19). Following alignment, the sequence variants [single-nucleotide variants (SNVs) and insertions or deletions (indels)] detected in the generated BAM files were assembled into a vcf format. The VariantStudio™ v2.1 analyser was used to perform variant filtering and annotation. The following criteria were used to define sequence variants—germline and somatic—and rule out mutation artefacts: (1) average wild-type read depth of > 500 × per pool (see Online Resource Table 2 in the electronic supplementary material), (2) occurrence in both forward and reverse sequencing pools, and (3) > 3% mutant allele frequency (MAF) in the merged vcf files. The Single Nucleotide Polymorphism Database (dbSNP) reference was used to separate germline from somatic sequence variants.

To assess the intra-assay variability of the NGS platform, we performed short-term precision assay by testing one sample in eight replicates in the same run. The inter-assay variability was assessed with the long-term precision assay by testing the same sample in three different runs. For each precision assay, we determined the coefficient of variation (CV).

2.4 QMC-PCR and HRM Analysis

As a means of validating the mutations detected by NGS, the samples were also analysed using the quick-multiplex-consensus (QMC)-PCR in conjunction with a high resolution melting (HRM) protocol as previously described [13]. Derivative and difference plots were generated to separate mutant from wild-type samples, as described elsewhere [13, 14].

2.5 Molecular Similarity Between Bx and Rx

To verify if the Bx were representative of the Rx at the molecular level, we investigated the similarities between the diagnostic Bx and Rx pairs by using three indices which have shown relevance in the clinical and biological behaviours of cancers: *somatic mutation profiles*, *MAF ratios*, and *allelic imbalance* (AI) status—within the limitations of the TruSight tumour targeted panel. Since each of the pairs of Bx and Rx are from the same tumours, they must be similar at the molecular level, i.e. not only must their mutation profiles match, but their mutant frequency ratios and AI scores must be in the same ranges.

A crude percentage concordance was used to calculate the extent to which the diagnostic Bx match the somatic mutation profiles, MAF ratios, and AI status of their corresponding Rx, whilst the kappa test [QuickCalcs (http://www.graphpad.com/quickcalcs/kappa2/) and Kappa (http://www.vassarstats.net/kappa.html)] was used to validate the crude percentage concordance test results [15, 16]. Mean difference in MAF between Rx and Bx was calculated using the online GraphPad software (http://www.graphpad.com).

2.6 Performance Evaluation of NGS-Based Somatic Mutation Profiling of Bx

As the 26-gene TruSight Tumour Somatic Mutation panel has translated into clinical use (http://www.clinicallabs.com.au/doctor/specialists-services/haematology-oncology/), we tested the following performance indices of the NGS-based somatic mutation profiling of Bx: sensitivity, specificity, and negative and positive predictive values (NPV and PPV). See Online Resource Table 3 in the electronic supplementary material. The performance indices as used here are merely to show the similarities between Rx and Bx at the molecular level and not strictly as diagnostic tests of accuracy.

3 Results

The NGS 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. The mean sequencing depth obtained was 14,803 (range 1366–44,577), whilst the limit of detection of the mutant alleles was 3%.

3.1 Paired Bx and Rx Mutation Profiles

A total of 78 and 81 somatic mutations were found in the Bx and Rx samples, respectively. Only two out of 25 tumour pairs (8%) displayed a wild-type genotype across all 26 genes included in the panel. The distribution of mutations detected in the 25 paired samples is shown in Table 1 and Online Resource Table 2. In sample 9, the *GNAS* c.2531G>A mutation was not detected in the Bx sample. In sample 13, only the Rx contained the *GNAS* c.2543C>T mutation. In sample 20, both the Bx and Rx contained the *TP53* c.524G>A mutation, but only the Rx contained the *TP53* c.23C>T mutation. Only eight out of 25 tumours (32%) contained the full complement of the *APC/KRAS(BRAF)/TP53* mutations of the Fearon and Vogelstein pathway. Furthermore, the frequency of *APC* mutations (56%) was lower than that of *TP53* mutations (68%), and this is consistent with published data. Although overall, the MAF was 1.003-fold lower in Rx than in Bx, on a mutation-by-mutation basis, the MAF showed no consistent pattern of abundance between the Rx and Bx samples. Moreover, there was no significant difference in the mean MAF between Rx and Bx samples (difference in mean MAF = 0.753, $P = 0.748$). Furthermore, the three mutations not detected in Bx were present in the matched

Table 1 Concordance by total mutant allele frequency (MAF) ratios

Sample no.	Gene	Mutation	Amino acid changes	Rx-MAF (%)	Bx-MAF (%)	Rx-MAF ratios	Bx-MAF ratios	Agreement
Rx&Bx1	*APC*	c.3997delA	Frameshift	8.96	3.94	1	1	
	APC	c.4216C>T	Q>[a]	11.94	4.07	1.332589286	1.032994924	C
	NRAS	c.35G>A	G>D	20.56	8.12	2.294642857	2.060913706	C
	TP53	c.273G>A	W>[a]	28.68	4.79	3.200892857	1.215736041	C
Rx&Bx2	*APC*	c.4216C>T	Q>[a]	21.99	12.96	1	1	
	KRAS	c.34G>T	G>C	33.5	22.45	1.523419736	1.732253086	C
Rx&Bx3	*APC*	c.4382_4383delAA		30.1	29.65	1	1	
	KRAS	c.35G>A	G>D	61.15	61.25	2.031561462	2.065767285	C
	TP53	c.428T>A	V>E	50.79	49.97	1.687375415	1.685328836	C
	FBXW7	c.1394G>A	R>H	32.07	30.91	1.065448505	1.042495784	C
	SMAD4	c.1609G>T	D>Y	52.87	54.63	1.756478405	1.842495784	C
Rx&Bx4	*APC*	c.4375_4376insC	Frameshift	12.91	12.09	1	1	
	KRAS	c.35G>T	G>V	35.48	33.59	2.748257165	2.778329198	C
	TP53	402_403delTT	Frameshift	45.2	45.68	3.50116189	3.778329198	C
Rx&Bx5	*TP53*	c.994-2A>C	Splice_variant	31.11	35.74	1	1	
Rx&Bx6	*APC*	c.4326delT	Frameshift	13.39	6.02	1	1	
	KRAS	c.35G>T	G>V	24.49	11.85	1.828976848	1.968438538	C
	PIK3CA	c.1633G>A	E>k	12.54	4.79	0.936519791	0.795681063	C
Rx&Bx7	*APC*	c.4732delT	Frameshift	11.66	9.56	1	1	
	KRAS	c.436G>A	A>T	11.83	9.45	1.01457976	0.988493724	C
Rx&Bx8	–	–		–	–			
Rx&Bx9	*APC*	c.4660_4661insA	Frameshift	25.1	14.91	1	1	
	KRAS	c.35G>T	G>V	29.5	15.94	1.175298805	1.069081154	C
	TP53	c.524G>A	R>H	13.53	20.14	0.539043825	1.350771294	D
	TP53	c.886C>T	R>[a]	3.9	3.8	0.155378486	0.254862508	C
	PIK3CA	c.331_333delAAG	Frameshift	10.23	17.37	0.407569721	1.16498994	D
	GNAS	c.2543C>T	R>H	3.38		0.134661355	0	D
Rx&Bx10	*APC*	c.3925G>T	E>[a]	5.53	15.2	1	1	
	APC	c.3940_3941delAG	Frameshift	6.66	16.69	1.204339964	1.098026316	C
	APC	c.3946C>T	A>V	6.83	16.96	1.235081374	1.115789474	C
	KRAS	c.38G>A	G>D	6.95	16.72	1.256781193	1.1	C
	TP53	c.623A>G	D>A	5.8	18.01	1.048824593	1.184868421	C
Rx&Bx11	*APC*	c.4216C>T	Q>[a]	19.94	27.48	1	1	
	KRAS	c.34G>T	G>C	18.73	24.58	0.939317954	0.894468705	C
	PIK3CA	c.290C>A	P>H	6.64	8.27	0.332998997	0.300946143	C
Rx&Bx12	*KRAS*	c.35G>A	G>D	21.99	40.36	1	1	
	TP53	c.523C>T	R>C	12.85	30.78	0.584356526	0.762636274	C
	FBXW7	c.1177C>T	R>[a]	12.49	27.86	0.567985448	0.690287413	C
		c.1513C>T	R>C	13.89	27.81	0.63165075	0.689048563	C
Rx&Bx13	*KRAS*	c.35G>T	G>V	10.15	4.82	1	1	
	TP53	c.186_193delAGC TCCCA	Frameshift	7.38	4.6	0.727093596	0.954356846	C
	PIK3CA	c.247_249invTTT	Frameshift	6.59	3.77	0.649261084	0.782157676	C
	GNAS	c.2531G>A	S>F	3.05		0.300492611	0	D
Rx&Bx14	*APC*	c.4385_4386delAG	Frameshift	6.68	12.83	1	1	
	PIK3CA	c.1637A>T	Q>L	5.46	11.28	0.817365269	0.8791894	C
	FBXW7	c.1136A>T	H>L	6.05	11.74	0.905688623	0.915042868	C
	PTEN	c.801+1G>A	Splice variant	10.1	20.73	1.511976048	1.615744349	C
	SMAD4	c.1082G>A	R>H	5.47	3.4	0.818862275	0.265003897	C
Rx&Bx15	–	–		–	–			

Table 1 (continued)

Sample no.	Gene	Mutation	Amino acid changes	Rx-MAF (%)	Bx-MAF (%)	Rx-MAF ratios	Bx-MAF ratios	Agreement
Rx&Bx16	*BRAF*	c.1780G>A	D>N	30.69	35.29	1	1	
	TP53	c.817C>T	R>C	50.18	60.3	1.63506028	1.708699348	C
Rx&Bx17	*APC*	c.4011_4012del GCinsTT	LQ>L[a]	24.32	29.75	1	1	
	KRAS	c.38G>A	G>D	13.38	16.33	0.550164474	0.548907563	C
	TP53	c.817C>T	R>C	34.63	41.96	1.423930921	1.410420168	C
	TP53	c.874A>G	K>E	13.03	15.4	0.535773026	0.517647059	C
	PIK3CA	c.1633G>A	E>K	3.44	3.93	0.141447368	0.13210084	C
	FBXW7	c.2065C>T	R>W	44.09	38.08	1.812911184	1.28	C
Rx&Bx18	*APC*	c.4529delG	Frameshift	45.82	14.88	1	1	
	APC	c.4530C>A	S>R	46.92	15.35	1.024006984	1.031586022	C
	KRAS	c.436G>A	A>T	51.03	28.2	1.113705805	1.89516129	C
	PIK3CA	c.316G>C	G>R	16.81	22.09	0.366870362	1.484543011	D
	FBXW7	c.1513C>T	R>C	48.5	17.97	1.058489742	1.20766129	C
Rx&Bx19	*BRAF*	c.1799T>A	V>E	19.69	27.53	1	1	
	TP53	c.404G>T	C>F	28.81	48.91	1.463179279	1.776607337	C
Rx&Bx20	*TP53*	c.23C>T	P>L	3.22		1		
	TP53	c.524G>A	R>H	18.71	22.23	5.810559006	–	D
Rx&Bx21	*APC*	c.4263_4264insA	Frameshift	16.68	16.34	1	1	
	APC	c.4264_4271del GAT CTTCC	Frameshift	16.57	16.18	0.993405276	0.990208078	C
	KRAS	c.35G>T	G>V	16.97	18.25	1.017386091	1.116891065	C
	TP53	c.874A>G	K>E	51.66	52.22	3.097122302	3.195838433	C
	PIK3CA	c.1633G>A	E>K	17.7	16.36	1.061151079	1.00122399	C
	SMAD4	c.1091T>G	L>W	10.93	12.35	0.655275779	0.755813953	C
		c.1094G>A	G>D	9.95	6.98	0.596522782	0.427172583	C
Rx&Bx22	*TP53*	c.524G>A	R>H	21.52	30.03			
Rx&Bx23	*APC*	c.2663C>T	A>V	13.63	24.05	1	1	
	APC	c.4222G>T	E>[a]	27.69	48.93	2.031548056	2.034511435	C
	KRAS	c.35G>A	G>D	25.38	41.69	1.862068966	1.733471933	C
	TP53	c.229C>T	P>S	51.04	47.78	3.744680851	1.986694387	C
	PIK3CA	c.325_327delGAA	Frameshift	13.75	26.56	1.008804109	1.104365904	C
	FBXW7	c.1393C>T	R>C	10.51	18.8	0.771093177	0.781704782	C
Rx&Bx24	*TP53*	c.844C>T	R>W	15.29	4.85	1	1	
	PTEN	c.795delA	Frameshift	3.15	1.85	0.206017005	0.381443299	C
Rx&Bx25	*TP53*	c.659A>G	Y>C	22.2	36.21			

Bx biopsy, *Rx* resection specimen

[a] *C* concordant, *D* discordant

Rx at frequencies of < 4%. There were no mutations in the Bx that were not seen in the Rx (Table 1). In all, only ten of the 26 genes in the TruSight panel were found to be mutated in the Rx and Bx samples.

3.2 Validation of Mutations

QMC-PCR in conjunction with HRM was used to validate the mutations identified, and initially 77 out of the 81 mutations (95.1%) were successfully validated (Online Resource Figure 1, see the electronic supplementary material). The remaining four mutations (4.9%) were only validated by HRM following minor allele enrichment by the modified co-amplification at lower denaturation temperature (COLD)-PCR protocol (Online Resource Figure 2). These four 'false negative' samples were subsequently reassigned as 'true positives'.

3.3 Allelic Imbalance

Quantification of heterozygous single-nucleotide polymorphisms (SNPs) was used to indicate allelic loss if there was

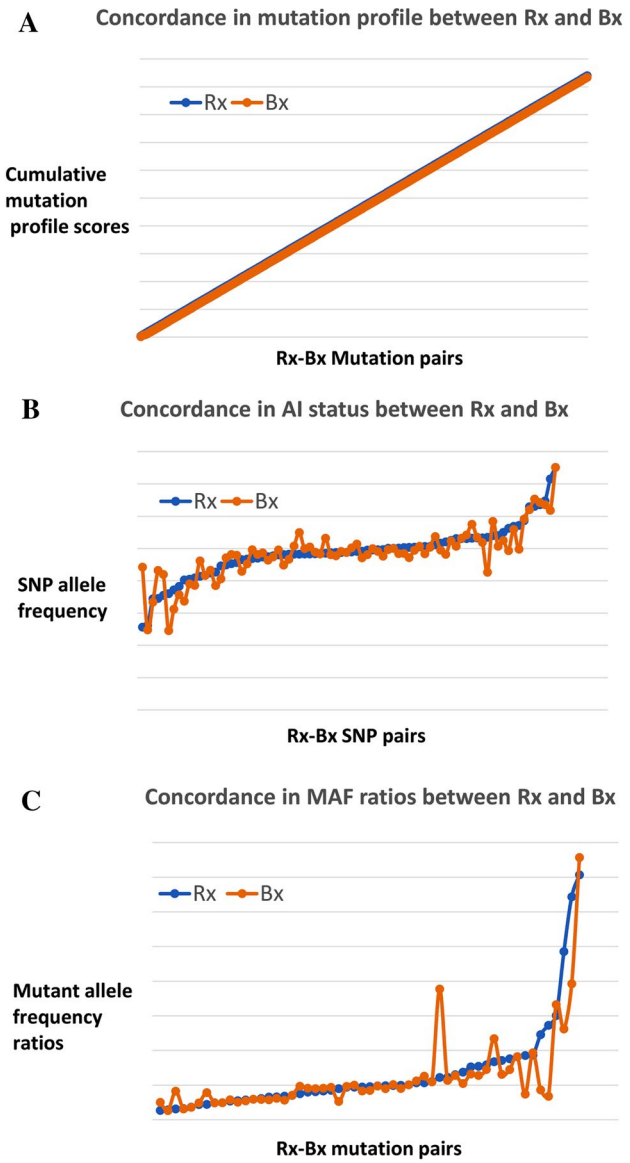

Fig. 1 Scatter plots showing the extent of agreement between Rx and Bx in the somatic mutation profile (**a**), AI status (**b**) and MAF ratios (**c**). All the detected mutations, regardless of the MAF, were included in the data that produced the somatic mutation profile and MAF ratio plots. The scatter plots show 'almost perfect' concordance in the somatic mutation profile to 'very good' and 'good' agreements in the AI status and MAF ratios, respectively. *AI* allelic imbalance, *Bx* biopsy, *MAF* mutant allele frequency, *Rx* resection specimen, *SNP* single-nucleotide polymorphism

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. Based on this, AI was found in Rx and matched Bx samples as shown in Table 2.

3.4 Concordance in Molecular Alteration Status Between Rx and Bx Pairs

To determine the similarity between Bx and their corresponding Rx at the molecular level, we determined the concordance in their somatic mutation profiles. A simple 'mutation-present-or-absent' count was used to determine the mutation status match between Bx and Rx. Only the ten mutated genes were used in this analysis, which included all 50 cases (25 Bx and 25 Rx). A total of 261 Rx–Bx mutation pairs were counted (Online Resource Figure 3). Of these, Bx and Rx showed concordance in 258 pairs (78 mutations and 180 no-mutations) and discordance in three pairs (all Rx: mutations/Bx: no-mutations). There was no Rx: no-mutation/ Bx: mutation pair. Also, all the mutations that matched were of the same bases in the same gene loci in Rx and Bx (Table 1). A crude percentage concordance of 98.85% (258/261) was calculated for the mutation status of Rx and Bx. The event indices were input into the online kappa calculators, QuickCalcs and Kappa. The result showed a Kappa of 0.971 [standard error (SE) of 0.016 and 95% confidence interval (CI) of 0.942–1.000], which is classified as 'almost perfect' agreement (see [17]) or 'very good' agreement (see Fig. 1a). Furthermore, the level of agreement between Bx and Rx in AI status was investigated. All 25 sample pairs with 80 informative SNP loci, cumulatively, were included in the analysis. AI status was categorised into three classes: AI with loss of wild-type allele (LWA, SNP % > 56.5%), AI with loss of polymorphic allele (LPA, SNP % < 43.3%) and nil AI (NAI, SNP % within normal range of 43.3% and 56.5%). The Rx/Bx pairs were scored concordant when their SNP classes matched, otherwise they were considered discordant. A total of 80 pairs were counted, comprising 51 out of 80 NAI pairs, seven out of 80 LWA pairs and 13 out of 80 LPA pairs. Discordance was found between Rx and Bx in ten out of 80 events (Rx/Bx: NAI/LPA = 3; NAI/LWA = 2; LPA/NAI = 2; LWA/NAI = 2; LPA/LWA = 0 and LWA/LPA = 0) (see Table 2). A crude percentage concordance of 88.75% (71/80) was calculated, giving a very good agreement between Bx and Rx for AI status (Fig. 1b). Kappa test also showed a 0.76 concordance (SE of 0.076 and 95% CI between 0.612 and 0.908).

Moreover, the total MAF ratios were compared between Rx and Bx. We reasoned that if Bx were truly representative of Rx, there should be some retention of the relative MAF ratios across the tumour body, despite the presence of clonal heterogeneity. A total of 20 out of 25 sample pairs, including only Rx/Bx pairs with two or more mutations in at least one of the Rx/Bx pairs, were included in this analysis. The MAF ratios for both Rx and Bx were calculated relative to the MAF of the first gene loci MAF in each Rx sample in Table 1. The Rx/Bx pair was considered concordant if both MAF ratios were either < 1 or > 1. If the MAF

Table 2 Allelic imbalance (AI) status of Rx and Bx

Sample pair no.	Informative SNP loci	Rx	Bx	AI status pair	Agreement per SNP locus
1	rs2228230	45.29	48.16	NAI/NAI	C
	rs1050171	50.64	50.39	NAI/NAI	C
2	rs3733542	49.67	47.62	NAI/NAI	C
	rs41115	48.31	50.48	NAI/NAI	C
	rs1050171	71.56	61.82	LWA/LWA	C
3	rs41115	48.24	50.73	NAI/NAI	C
	rs1050171	50.62	48.36	NAI/NAI	C
	rs2023748	46.73	45.25	NAI/NAI	C
	rs41737	52.49	52.3	NAI/NAI	C
	rs1137282	26.2	24.76	LPA/LPA	C
4	rs1050171	56.24	49.35	NAI/NAI	C
	rs2023748	40.86	38.53	LPA/LPA	C
	rs41737	44.63	40.59	NAI/LPA	D
5	rs2228230	34.51	43.18	LPA/LPA	C
	rs1042522	25.62	44.24	LPA/NAI	D
6	rs41115	37.2	31.16	LPA/LPA	C
7	rs2228230	42.42	43.24	LPA/LPA	C
	rs41115	62.91	62	LWA/LWA	C
	rs2229066, rs17290559	48.83	48.74	NAI/NAI	C
	rs2023748	44.9	47.1	NAI/NAI	C
	rs41737	50.31	48.49	NAI/NAI	C
	rs1137282	40.46	39.08	LPA/LPA	C
	rs1042522	63.53	64.22	LWA/LWA	C
8	rs41115	64.73	63.53	LWA/LWA	C
	rs1137282	47.31	48.64	NAI/NAI	C
	rs1042522	53.95	58.39	NAI/LWA	D
9	rs2228230	34.4	33.38	LPA/LPA	C
10	rs3733542	50.4	49.45	NAI/NAI	C
	rs41115	50.37	47.18	NAI/NAI	C
11	rs2023748	56.85	55.89	LWA/NAI	D
	rs41737	58.6	59.13	LWA/LWA	C
	rs1042522	75.14	75.12	LWA/LWA	C
12	rs1050171	63.04	65.31	LWA/LWA	C
	rs41737	42.69	38.51	LPA/LPA	C
	rs1042522	40.24	33.67	LPA/LPA	C
13	rs41737	57.12	49.75	LWA/NAI	D
14	rs2228230	47.47	46.37	NAI/NAI	C
	rs41115	47.75	47.38	NAI/NAI	C
	rs1050171	53.44	42.62	NAI/LPA	D
	rs41737	46.97	48.12	NAI/NAI	C
15	rs3733542	48.36	48.86	NAI/NAI	C
	rs41115	54.99	52.48	NAI/NAI	C
	rs2229066, rs17290559	41.66	41.61	LPA/LPA	C
	rs2023748	49.58	49.1	NAI/NAI	C
16	rs2228230	48.77	47.72	NAI/NAI	C
	rs41115	48.42	48.31	NAI/NAI	C
17	rs2228230	48.03	49.54	NAI/NAI	C
	rs3733542	50.1	48.41	NAI/NAI	C

Table 2 (continued)

Sample pair no.	Informative SNP loci	Rx	Bx	AI status pair	Agreement per SNP locus
	rs41115	53.2	57.51	NAI/LWA	D
	rs1050171	48.24	54.96	NAI/NAI	C
18	rs2228230	48.78	49.11	NAI/NAI	C
	rs3733542	48.9	50.2	NAI/NAI	C
	rs41115	38.28	35.64	LPA/LPA	C
	rs1050171	52.07	48.21	NAI/NAI	C
	rs1137282	53.27	53.53	NAI/NAI	C
	rs1042522	51.12	53.7	NAI/NAI	C
19	rs41115	49.12	51.38	NAI/NAI	C
20	rs3733542	50.07	49.69	NAI/NAI	C
	rs41115	53.99	50.71	NAI/NAI	C
	rs1050171	35.47	41.93	LPA/LPA	C
	rs2023748	41.4	46.13	LPA/NAI	D
21	rs1050171	49.13	47.1	NAI/NAI	C
	rs35775721	48.11	44.83	NAI/NAI	C
	rs33917957	46.52	42.93	NAI/LPA	D
22	rs56391007	53.05	50.68	NAI/NAI	C
	NM_001127500.1	49.54	49.94	NAI/NAI	C
	NM_001127500.1	45.63	47.9	NAI/NAI	C
	NM_001127500.1	53.3	51.89	NAI/NAI	C
	NM_033360.2	48.23	46.66	NAI/NAI	C
	rs1042522	53.09	54.2	NAI/NAI	C
23	rs55789615	50.07	50.03	NAI/NAI	C
	rs1050171	51.67	49.47	NAI/NAI	C
24	rs3733542	49.49	48.09	NAI/NAI	C
	rs41115	36.03	24.57	LPA/LPA	C
	rs1050171	50.49	50.7	NAI/NAI	C
	rs2023748	46.86	49.58	NAI/NAI	C
	rs41737	48.55	53.15	NAI/NAI	C
25	rs41115	48.3	49.92	NAI/NAI	C
	rs1050171	48.7	48.05	NAI/NAI	C
	rs1137282	53.07	53.17	NAI/NAI	C

Bx biopsy, *C* concordant, *D* discordant, *LPA* allelic imbalance with loss of polymorphic allele, *LWA* allelic imbalance with loss of wild-type allele, *NAI* nil allelic imbalance, *Rx* resection specimen, *SNP* single-nucleotide polymorphism

ratios for the Bx/Rx pair were < 1 and > 1, but were within 1 ± 0.05, they were also considered concordant. Otherwise, they were taken as discordant. Also, samples in which one member of the pair was missing a corresponding mutation were considered discordant and were classed as being in the $Bx < 1/Rx > 1$ category as the Rx MAF ratios in all those cases were > 1. A total of 58 mutation pairs were counted, comprising 52 concordant observations between Rx and Bx (comprising 45 MAF ratio pairs < 1, six MAF ratio pairs > 1 and one MAF ratio pair $= 1 \pm 0.05$) and six discordant observations (all $Bx < 1/Rx > 1$). There were zero $Bx > 1/Rx < 1$ MAF ratio pairs. A crude percentage concordance rate of 89.6% was calculated for the total MAF ratios of Rx and Bx. Kappa was 0.651 (SE = 0.128, 95% CI 0.400–0.901). Both tests again returned 'good' to 'very good' agreement scores between the MAF ratios of Bx and Rx samples (Fig. 1c).

3.5 Performance Evaluation of NGS-Based Somatic Mutation Profiling of Bx

We evaluated the use of Bx for mutation detection by NGS using established tests of performance (Online Resource Table 3). Using the Rx as the 'gold standard' samples and taking each of the somatic mutations detected (or not detected) as individual observations, the following parameters were derived for Bx samples: number of true positive

tests = 78, true negatives = 180, false positives = 0 and false negatives = 3.

The indices of performance obtained for Bx include sensitivity of 96.3% with a false negative rate of 3.7%, specificity of 100% with a false positive rate of 0%, a PPV (precision) value of 100%, NPV of 98.4%, accuracy of 98.85%, and a false discovery rate of 0%, altogether indicating a high performance of Bx as suitable samples for molecular testing by NGS.

4 Discussion

Recent advances towards personalised medicine are driven by the identification of targetable mutations. For example, treatment of non-small cell lung cancer patients with gefitinib is dependent upon *EGFR* mutation status [18]. Herceptin administration is only considered in a subset of breast and gastric cancer patients with *HER2* amplification [18, 19]. In CRC patients with advanced disease, mutation screening of *KRAS* and *NRAS* is required prospectively if anti-EGFR monoclonal antibody therapies are being considered, as responses have only been seen in wild-type tumours [5, 20].

Where targeted neoadjuvant chemotherapy is being offered to patients, mutation screening must be carried out on the diagnostic Bx specimen. Thus, the question arises as to whether a Bx specimen, which represents a tiny proportion of the tumour, is adequately representative of the whole tumour and thus can be used in patient stratification. Previously, we and others showed that FFPE diagnostic Bx tissues were adequate for testing microsatellite instability and other molecular alterations in CRC by low throughput methods such as HRM analysis, direct sequencing, pyrosequencing, and Therascreen Amplification Refractory Mutation System (ARMS)-Scorpion [11, 21]. Furthermore, other groups have demonstrated the feasibility and reliability of the use of small diagnostic Bx for molecular testing by NGS [22–25]. In this study, despite the use of low-quality DNA template derived from FFPE tissue, we obtained a mean sequencing depth of 14,803 (range 1366–44,577) and the limit of detection for the mutant alleles was 3%. There was good short-term and long-term precision, and all 81 somatic mutations detected using the TruSight panel were also validated by QMC-PCR and HRM. Validation of low-level mutations required COLD-PCR to further enrich the mutant allele population.

In our sample set, the frequency of detected gene mutations was within the range of previously published literature [26, 27]. The most frequent mutations were in *TP53*, whilst *APC* mutation was found in 56% of tumours. The sensitivity of targeted NGS analysis allowed the detection, in the Bx samples, of all but three of the 81 mutations detected in the paired Rx. There was no significant difference in the mean MAF between Rx and Bx samples.

More importantly, we compared the degree of similarity between the Rx and Bx pairs at the molecular level using well established statistical tests and markers which have been shown to have biological and clinical importance [5, 17, 28–33]. The presence-or-absence-of-mutation-type and the AI status tests showed very good concordance between the Rx and the Bx samples, an indication that the latter were adequately representative of the former. Furthermore, we applied the MAF ratios to test the degree of similarity between the two Bx types and found a 'good' to 'very good' concordance between them. Whilst somatic mutation profiles and AI status have established biological, prognostic and predictive utilities, MAF is currently under active clinical research for use as a marker for the estimation of tumour heterogeneity and prediction of cancer survival, targeted therapy response and the risks and foci of tumour metastases [29–33].

Furthermore, the Bx samples showed relatively high indices of performance as potential clinical test materials for somatic mutation detection by NGS, an indication that Bx is an adequate material for molecular testing for neoadjuvant therapy.

Although our data indicate that Bx specimens represent a feasible material for molecular testing, to increase the probability of sampling of the dominant clone, some factors should be considered when interpreting data from tumour Bx specimens. For example, from where was the tissue taken? The centre or invasive edge of the tumour? A study performed by Baldus et al. [34] demonstrated a discrepancy in the frequency of mutations in *KRAS*, *BRAF* and *PIK3CA* by 8%, 1% and 5%, respectively, between the centre and the invasive edge of colorectal tumours [34], with one explanation of this discrepancy being that the invasive edges are probably more prone to stromal contamination than the central portions of the tumour. Another factor is related to tumour clonal heterogeneity [35]. Although we did find overall a strong agreement between Rx and Bx at the molecular level, we observed that a proportion of the Rx and Bx showed MAF discrepancies at some loci and that three out of 81 Bx samples did not show the corresponding mutations which were observed in the Rx samples with MAFs < 4%. On the basis of these factors, we advocate that diagnostic Bx with intent for molecular testing should sample multiple tumour areas to enhance mutation detection.

This study is limited by the number of SNPs that could be interrogated to allow a more comprehensive AI status analysis—the TruSight panel targets gene exons which have lower SNP densities compared to introns. Another limitation of this study is the small sample size used for the evaluation of Bx as a suitable candidate for molecular testing by NGS. The use of a larger sample size is perhaps necessary

to validate the use of diagnostic Bx as an adequate Bx for mutation detection on the NGS platform.

In conclusion, we have shown a high concordance between matched Bx and Rx within the mutation distributions of the genes in the TruSight tumour panel, suggesting that the use of diagnostic Bx is not only feasible, but also representative of the entire tumour, and thus can be used for predictive mutation screening.

Acknowledgements The authors would like to thank Gareth Cross for enabling the process of NGS data generation and analyses.

Compliance with Ethical Standards

Funding This work was funded by Universities of Nottingham (for MI) and Leeds (for SDR).

Conflict of interest All the authors (HH-K, HOE, KM, WF, NSA, SDR and MI) declare that they have no conflicts of interest in publishing this manuscript.

Ethical approval and informed consent 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).

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136:E359–86. https://doi.org/10.1002/ijc.29210.
2. Cancer Research UK. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/. Accessed 20 June 2018.
3. LeBlanc VG, Marra MA. Next-generation sequencing approaches in cancer: where have they brought us and where will they take us? Farah CS, Cho WC, eds. Cancers. 2015;7(3):1925–58. https://doi.org/10.3390/cancers7030869.
4. Li J, Wang T, Zhang X, Yang X. The contribution of next generation sequencing technologies to epigenome research of stem cell and tumorigenesis. Hum Genet Embryol. 2011;S2:001. https://doi.org/10.4172/2161-0436.S2-001.
5. Schmiegel W, Scott RJ, Dooley S, Lewis W, Meldrum CJ, Pockney P, et al. Blood-based detection of RAS mutations to guide anti-EGFR therapy in colorectal cancer patients: concordance of results from circulating tumor DNA and tissue-based RAS testing. Mol Oncol. 2017;11(2):208–19.
6. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, et al. Molecular biomarkers for the evaluation of colorectal cancer: guideline from the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. J Mol Diagn. 2017;19(2):187–225.
7. Santos MD, Silva C, Rocha A, Nogueira C, Castro-Poças F, Araujo A, et al. Predictive clinical model of tumor response after chemoradiation in rectal cancer. Oncotarget. 2017;8(35):58133–51.
8. Lim SH, Chua W, Henderson C, Ng W, Shin JS, Chantrill L, et al. Predictive and prognostic biomarkers for neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Crit Rev Oncol Hematol. 2015;96(1):67–80.
9. FoxTROT Collaboration Group. Feasibility of preoperative chemotherapy for locally advanced, operable colon cancer: the pilot phase of a randomised controlled trial. Lancet Oncol. 2012;13(11):1152–60.
10. Jakobsen A, Andersen F, Fischer A, Jensen LH, Jørgensen JCR, Larsen O, et al. Neoadjuvant chemotherapy in locally advanced colon cancer. A phase II trial. Acta Oncol. 2015;54(10):1747–53.
11. Fadhil W, Ibrahem S, Seth R, AbuAli G, Ragunath K, Kaye P, et al. The utility of diagnostic biopsy specimens for predictive molecular testing in colorectal cancer. Histopathology. 2012;61(6):1117–24.
12. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306–13.
13. Fadhil W, Ibrahem S, Seth R, Ilyas M. 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. 2010;63(2):134–40.
14. Seth R, Crook S, Ibrahem S, Fadhil W, Jackson D, Ilyas M. Concomitant mutations and splice variants in *KRAS* and *BRAF* demonstrate complex perturbation of the Ras/Raf signalling pathway in advanced colorectal cancer. Gut. 2009;58(9):1234–41.
15. GraphPad QuickCalcs. http://www.graphpad.com/quickcalcs/kappa2/. Accessed 20 June 2018.
16. Lowry R. VassarStats: website for statistical computation. http://vassarstats.net/Kappa. Accessed 20 June 2018.
17. Viera AJ, Garrett JM. Understanding inter-observer agreement: the kappa statistic. Fam Med. 2005;37(5):360–3.
18. Gajria D, Chandarlapaty S. HER2-amplified breast cancer: mechanisms of trastuzumab resistance and novel targeted therapies. Expert Rev Anticancer Ther. 2011;11(2):263–75.
19. Abrahao-Machado LF, Scapulatempo-Neto C. HER2 testing in gastric cancer: an update. World J Gastroenterol. 2016;22(19):4619–25.
20. Bokemeyer C, Cutsem EV, Rougier P, Ciardiello F, Heeger S, Schlichting M, et al. Addition of cetuximab to chemotherapy as first-line treatment for KRAS wild-type metastatic colorectal cancer: pooled analysis of the CRYSTAL and OPUS randomised clinical trials. Eur J Cancer. 2012;48(10):1466–75.
21. Krol LC, 't Hart NA, Methorst N, Knol AJ, Prinsen C, Boers JE. Concordance in KRAS and BRAF mutations in endoscopic biopsy samples and resection specimens of colorectal adenocarcinoma. Eur J Cancer. 2012;48(7):1108–15.
22. Ku BM, Heo MH, Kim J-H, Cho BC, Cho EK, Min YJ, et al. Molecular screening of small biopsy samples using next-generation sequencing in Korean patients with advanced non-small cell lung cancer: Korean Lung Cancer Consortium (KLCC-13-01). J Pathol Transl Med. 2018;52(3):148–56.
23. Zheng G, Tsai H, Tseng L-H, Gocke CD, Eshleman JR, Netto G, Lin M-T. Test feasibility of next-generation sequencing assays in clinical mutation detection of small biopsy and fine needle aspiration specimens. Am J Clin Pathol. 2016;145:696–702.
24. Illei PB, Belchis D, Tseng L-H, Nguyen D, De Marchi F, Haley L, et al. Clinical mutational profiling of 1006 lung cancers by next generation sequencing. Oncotarget. 2017;8(57):96684–96.
25. DiBardino D, Rawson D, Saqi A, Heymann J, Pagan C, Bulman W. Next-generation sequencing of non-small cell lung cancer using a customized, targeted sequencing panel: emphasis on small biopsy and cytology. CytoJournal 2017;14(1):7. http://www.cytojournal.com/text.asp?2017/14/1/7/202602. Accessed 20 June 2018.
26. Catalogue of somatic mutation in cancer at https://cancer.sanger.ac.uk/. Accessed 20 June 2018.
27. The Cancer Genome Atlas at https://cancergenome.nih.gov/. Accessed 20 June 2018.

28. Nurwidya F, Takahashi F, Takahashi K. Gefitinib in the treatment of non-small cell lung cancer with activating epidermal growth factor receptor mutation. J Nat Sci Biol Med. 2016;7(2):119–23. https://doi.org/10.4103/0976-9668.184695.
29. Tejpar S, Bertagnolli M, Bosman F, Lenz H-J, Garraway L, Waldman F, Warren R, et al. Prognostic and predictive biomarkers in resected colon cancer: current status and future perspectives for integrating genomics into biomarker discovery. Oncologist. 2010;15(4):390–404.
30. Dienstmann R, Elez E, Argiles G, et al. Analysis of mutant allele fractions in driver genes in colorectal cancer—biological and clinical insights. Mol Oncol. 2017;11(9):1263–72. https://doi.org/10.1002/1878-0261.12099.
31. Shen S, Wei Y, Zhang R, et al. Mutant-allele fraction heterogeneity is associated with non-small cell lung cancer patient survival. Oncol Lett. 2018;15(1):795–802. https://doi.org/10.3892/ol.2017.7428.
32. Ono A, Kenmotsu H, Watanabe M, Serizawa M, Mori K, Imai H, et al. Mutant allele frequency predicts the efficacy of EGFR-TKIs in lung adenocarcinoma harboring the L858R mutation. Ann Oncol. 2014;25(10):1948–53. https://doi.org/10.1093/annonc/mdu251.
33. Doma V, Papp O, Rásó E, Barbai T, Reiniger L, Vízkeleti L, Timar J. Alteration of mutant allele frequency in visceral metastases of melanoma. J Clin Oncol. 2018;36(15_suppl):e21528.
34. Baldus SE, Schaefer K-L, Engers R, Hartleb D, Stoecklein NH, Gabbert HE. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. Clin Cancer Res. 2010;16(3):790–9.
35. Yun J, Rago C, Cheong I, Pagliarini R, Angenendt P, Rajagopalan H, et al. Glucose deprivation contributes to the development of *kras* pathway mutations in tumor cells. Science. 2009;325(5947):1555–9.