FULL PAPER

BJC

British Journal of Cancer (2017) 117, 203–209 | doi: 10.1038/bjc.2017.168

BJC OPEN

Keywords: genotype; PDL1; KIT

Molecular profiling of signet ring cell colorectal cancer provides a strong rationale for genomic targeted and immune checkpoint inhibitor therapies

Muhammad A Alvi[1], Maurice B Loughrey[1,2], Philip Dunne[1], Stephen McQuaid[1], Richard Turkington[3], Marc-Aurel Fuchs[1], Claire McGready[1], Victoria Bingham[1], Brendan Pang[4], Wendy Moore[1], Perry Maxwell[1], Mark Lawler[3], Jacqueline A James[1], Graeme I Murray[5], Richard H Wilson[3] and Manuel Salto-Tellez*[,1,2]

[1]*Northern Ireland Molecular Pathology Laboratory, Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast BT9 7AE, UK;* [2]*Department of Histopathology, Royal Victoria Hospital, Belfast Health and Social Care Trust, Belfast, UK;* [3]*Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, UK;* [4]*Department of Pathology, National University Hospital, National University Health System, Singapore, Singapore and* [5]*Department of Pathology, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen AB25 2ZD, UK*

Background: Signet ring cell colorectal cancer (SRCCa) has a bleak prognosis. Employing molecular pathology techniques we investigated the potential of precision medicine in this disease.

Methods: Using test ($n=26$) and validation ($n=18$) cohorts, analysis of mutations, DNA methylation and transcriptome was carried out. Microsatellite instability (MSI) status was established and immunohistochemistry (IHC) was used to test for adaptive immunity (CD3) and the immune checkpoint PDL1.

Results: DNA methylation data split the cohorts into hypermethylated ($n=18$, 41%) and hypomethylated groups ($n=26$, 59%). The hypermethylated group predominant in the proximal colon was enriched for CpG island methylator phenotype (CIMP), *BRAF V600E* mutation and MSI ($P<0.001$). These cases also had a high $CD3^+$ immune infiltrate ($P<0.001$) and expressed PDL1 ($P=0.03$ in intra-tumoural lymphoid cells). The hypomethylated group predominant in the distal colon did not show any characteristic molecular features. We also detected a common targetable *KIT* mutation (c.1621A>C) across both groups. No statistically significant difference in outcome was observed between the two groups.

Conclusions: Our data show that SRCCa phenotype comprises two distinct genotypes. The MSI^+/$CIMP^+$/*BRAF V600E*$^+$/$CD3^+$/$PDL1^+$ hypermethylated genotype is an ideal candidate for immune checkpoint inhibitor therapy. In addition, one fourth of SRCCa cases can potentially be targeted by *KIT* inhibitors.

Colorectal cancer (CRC) rates are on the decline in the US and Western Europe, but incidence of signet ring cell colorectal cancer (SRCCa) has remained steady (Gopalan *et al*, 2011; Arnold *et al*, 2017). These are highly malignant, dedifferentiated adenocarcinomas and comprise around 0.1–2.4% of all CRC cases (Anthony *et al*, 1996). Primary SRCCas are most often diagnosed at an advanced stage, and typically have a dismal prognosis with average five year survival rates of around 20% (Nitsche *et al*, 2013).

Molecular pathology of SRCCa is not well understood and it is unclear whether the signet ring cell phenotype carries a distinct

*Correspondence: Professor M Salto-Tellez; E-mail: m.salto-tellez@qub.ac.uk

Revised 6 May 2017; accepted 15 May 2017; published online 8 June 2017

genotype as well. Because of the rarity of this cancer most published studies are either case reports or retrospective epidemiological and clinicopathological analyses. High frequency of *BRAF* mutations, microsatellite instability (MSI) and CpG island methylator phenotype (CIMP) have been reported along with predominance in proximal colon and the female gender (Kakar *et al*, 2012). But to date there has not been any multi-omics study conducted to comprehensively study the molecular pathology of this disease. This is necessary for two reasons: (a) to shed light on whether SRCCa has a molecular profile distinct from other CRC subtypes, and (b) to identify novel biomarkers and therapeutic targets.

MATERIALS AND METHODS

Patient samples. For the test cohort, patients were identified from the pathology archives of Belfast Health and Social Care Trust (BHSCT) in Northern Ireland and formalin fixed paraffin-embedded (FFPE) tissue blocks were made available by the Northern Ireland Biobank. For the validation cohort, patients were both identified and FFPE tissue blocks made available from the Grampian Biorepository in Scotland. Ethical approval was provided by the Northern Ireland Biobank scientific access group committee (study number–NIB14-0139), the Grampian Biorepository scientific access group committee (tissue request–TR000058) and the NHS Health Research Authority North West–Preston research ethics committee (reference–15/NW/0855). No written consent was required from patients for the use of FFPE tissue blocks and anonymised demographic and clinicopathological data. All identified patients were reviewed by two expert pathologists (MST and MBL for test cohort/GIM and MBL for validation cohort). Only cases that fulfilled the WHO criteria of greater than 50% of the tumour comprising of signet ring cells were selected for the study (Bosman, 2010).

Nucleic acid extraction. Representative normal (furthest from the tumour) and tumour (highest cellularity of signet ring tumour cells) FFPE tissue blocks were selected after haematoxylin and eosin (H&E) slide review. The H&E slides were then annotated for the epithelial layer in normal blocks and signet ring cell rich areas in tumour blocks (MST and MBL). $5 \times 10\,\mu$m and $5 \times 8\,\mu$m sections were cut onto glass slides for DNA and RNA extraction, respectively. Annotated areas were macrodissected using sterile scalpel blades into 1.5 ml microcentrifuge tubes. DNA extraction was done using Maxwell 16 FFPE Plus LEV DNA Purification Kit (Promega, Southampton, UK), and RNA using the RNeasy FFPE Kit (Qiagen, Manchester, UK). Quantification was conducted using NanoDrop 2000 (Thermo Fisher Scientific Inc., Loughborough, UK) unless mentioned otherwise.

Next generation sequencing. Next generation sequencing (NGS) was performed on the entire test cohort tumour samples. TaqMan RNase P Detection Reagents Kit was used to quantify 10 ng of DNA and library prepared using the Ion AmpliSeq Library Kit 2.0 and Cancer Hotspot Panel v2 (Thermo Fisher Scientific Inc.). Sequencing was performed on the Ion PGM System according to manufacturer's instructions and our previously published protocols (McCourt *et al*, 2013; Alvi *et al*, 2015).

DNA methylation. DNA methylation arrays were performed on both test and validation cohort tumour samples and additionally on 10 randomly selected normal samples from the test cohort. We used the Infinium 450k arrays (Illumina Inc., Cambridge, UK) following the manufacturer's instructions and our previously published protocol (Alvi *et al*, 2015). Total 200 ng of DNA as quantified using Qubit Fluorometric Quantitation assay (Thermo Fisher Scientific Inc.) was used and arrays were scanned using iScan (Illumina Inc.).

Gene expression. Gene expression arrays were performed on test cohort tumour samples and 10 randomly selected normal samples. The Whole-Genome DASL HT assay was used in combination with the HumanHT-12 v4 BeadChip (Illumina Inc.) according to manufacturer's instructions and our previously published protocol (Alvi *et al*, 2015). Around 200 ng of total RNA was used as quantified by Qubit Fluorometric Quantitation assay and chips were scanned using iScan.

Sanger sequencing. Sequencing was carried out using the BigDye Terminator v3.1 Cycle Sequencing Kit on the ABI 3500 Genetic Analyzer (Thermo Fisher Scientific Inc.) using manufacturer's instructions. Primers were designed using NCBI primer designing tool with M13 overhangs. All PCRs were carried out using AmpliTaq Gold 360 Master Mix (Thermo Fisher Scientific Inc.), and cleaned using ExoSAP-IT (Affymetrix, UK). Approximately 10–50 ng of DNA was used for each reaction.

Microsatellite instability analysis. MSI status was evaluated using MSI Analysis System, Version 1.2 (Promega) according to the manufacturer's instructions. We tested five mononucleotide repeat markers (BAT-25, BAT-26, NR-21, NR-24 and MONO-27), which were co-amplified using fluorescently labelled primers and analysed on an ABI 3500 Genetic Analyzer. Approximately 10–50 ng of DNA was used for each reaction.

***BRAF* V600E mutation assay.** Cobas 4800 BRAF V600 mutation test kit (Roche Molecular Systems Inc., Burgess Hill, UK) was used to look for *BRAF V600E* mutation according to the manufacturer's instructions. Around 125 ng of DNA was used for each reaction.

Immunohistochemistry. PDL1 and CD3 immunohistochemistry was carried out on 3 μm full face sections using PD-L1/CD274 (SP142) antibody (Spring Bioscience, CA, USA) at 1 : 40 dilution and CONFIRM anti-CD3 (2GV6) rabbit monoclonal antibody (Ventana, UK) respectively. In addition OptiView amplification kit was used for PDL1 antibody. Staining was carried out on Ventana Benchmark XT platform with the OptiView Universal DAB Detection Kit (Ventana Medical Systems, Burgess Hill, UK).

PDL1 scoring was performed separately for peritumoural lymphoid follicles (PLF), intra-tumoural lymphoid cells (ILC) and tumour epithelial cells (TEC). Scoring criteria used was 0 (negative) for no cell staining and 1 (positive) for any number of cells staining. CD3 staining, assessed in ILCs only, was scored semi-quantitatively using a three tiered scoring system (1 for the lowest counts observed and 3 for the highest).

Data analysis. For NGS data vcf files were generated from the torrent server using the variantCaller plugin (Life Technologies, Loughborough, UK) and imported into Ion Reporter 5.0 (Thermo Fisher Scientific Inc.) for annotation. Methylation and gene expression array data was analysed using GenomeStudio methylation and expression modules version 1.9.0 respectively (Illumina Inc.). Sanger sequencing data were viewed and confirmed with Finch TV version 1.4.0 (Geospiza Inc., WA, USA). Gene set enrichment analysis (www.broadinstitute.org/gsea) was used for pathway analysis using default settings. Assignment of patient samples into their respective consensus molecular subtyping (CMS) groups based on gene expression data was carried out using the 'CMSclassifier' package in R version 3.2.4 (The R Foundation for Statistical Computing, Austria; Guinney *et al*, 2015). For comparing data between groups, using Prism version 5 (GraphPad Software, CA, USA) a *t*-test was performed for continuous variables and Fisher's exact test for categorical variables. Cox proportional hazards analysis to look for

associations between molecular and clinicopathological data were conducted using Stata version 11.2 (StataCorp, TX, USA).

RESULTS

Patient cohorts. Total of 26 and 18 patients were identified from the BHSCT and the Grampian Biorepository, respectively. We did not observe any statistically significant difference between the two cohorts in terms of demographics or clinicopathological data (Supplementary Table 1).

DNA methylation. From the test cohort, based on beta values, most variable probes were identified using a s.d. cut-off of 0.25. This generated a list of 875 probes. These probes were used for unsupervised hierarchical clustering using the manhattan metric and were able to split the 26 sample cohort into distinct hypermethylated ($n=9$) and hypomethylated ($n=17$) groups. The same probes were also able to split the validation cohort into hypermethylated ($n=9$) and hypomethylated ($n=9$) groups. As shown in Figure 1 only 300 (enclosed in red) out of the 875 probes are consistently differentially methylated between the two groups. The full list is available in Supplementary Table 2 and raw data can be obtained from GSE79740.

As shown in Supplementary Figure 1 the hypermethylated group was also CIMP positive. We were also able to identify genes consistently hypermethylated and hypomethylated across all tumour samples compared to normal tissue with potential as diagnostic biomarkers (Supplementary Table 3).

Mutations. As shown in Figure 1, compared to the hypomethylated group, we observed the hypermethylated group to be enriched for *BRAF V600E* mutation ($P<0.001$ in test, validation and both cohorts combined together). This was also confirmed using a PCR based assay. Other mutations were also observed in the test cohort using the 50 gene hotspot panel (PRJNA316428, Supplementary Table 4). According to the COSMIC database, compared to colorectal adenocarcinoma average we observed higher frequencies of *TP53* (69% *vs* 44%), *BRAF* (31% *vs* 13%) and *KIT* (34% *vs* 8%) mutations. At the same time a lower frequency of *APC* (35% *vs* 45%), *KRAS* (12% *vs* 34%), *PIK3CA* (4% *vs* 14%) and *ATM* (4% *vs* 18%) mutations was observed (Figure 2 and Supplementary Table 4). The number of mutant genes in each sample also varied ranging between 1 and 11 with an average of 2.7 mutant genes per sample (out of the 50 tested by panel). This average was 3.9 in the hypermethylated and 2.1 in the hypomethylated group ($P<0.05$).

The *KIT* mutations detected by NGS in the test cohort were similar in eight out of the nine cases (c.1621A>C). This was validated using a Sanger sequencing assay in both the test and validation cohorts and an additional three cases were found in the validation cohort (Forward primer: GTTGTAAAACGACGGCCAGUCGTAGCTGGCATGATGTGC, R primer: CACAGGAAACAGCTATGACCTCTGGAGAGAGAACAAATAAATGGT).

Gene expression. Gene set enrichment analysis was used to look at pathway enrichment at the gene expression level in the test cohort. Using the 'hallmark 50 gene sets' we identified 19 gene sets enriched in the hypermethylated group and 4 in the hypomethylated group ($q<0.05$). The top three were 'MTOR signalling', 'MYC targets V1' and 'E2F targets' in the hypermethyated group and 'epithelial mesenchymal transition', 'myogenesis' and 'apical junction' in the hypomethylated group (Supplementary Figure 2 and Supplementary Table 5).

Gene expression data (GSE79793) was also merged with DNA methylation data using GenomeStudio and spearman correlation coefficients were calculated for every combination of methylation and expression array probes (Supplementary Table 6). We observed 5725 combinations (2088 gene probes) with an inverse relationship in the hypermethylated group and only 753 combinations (439 gene probes) in the hypomethylated group highlighting the impact of differential methylation between the two (spearman

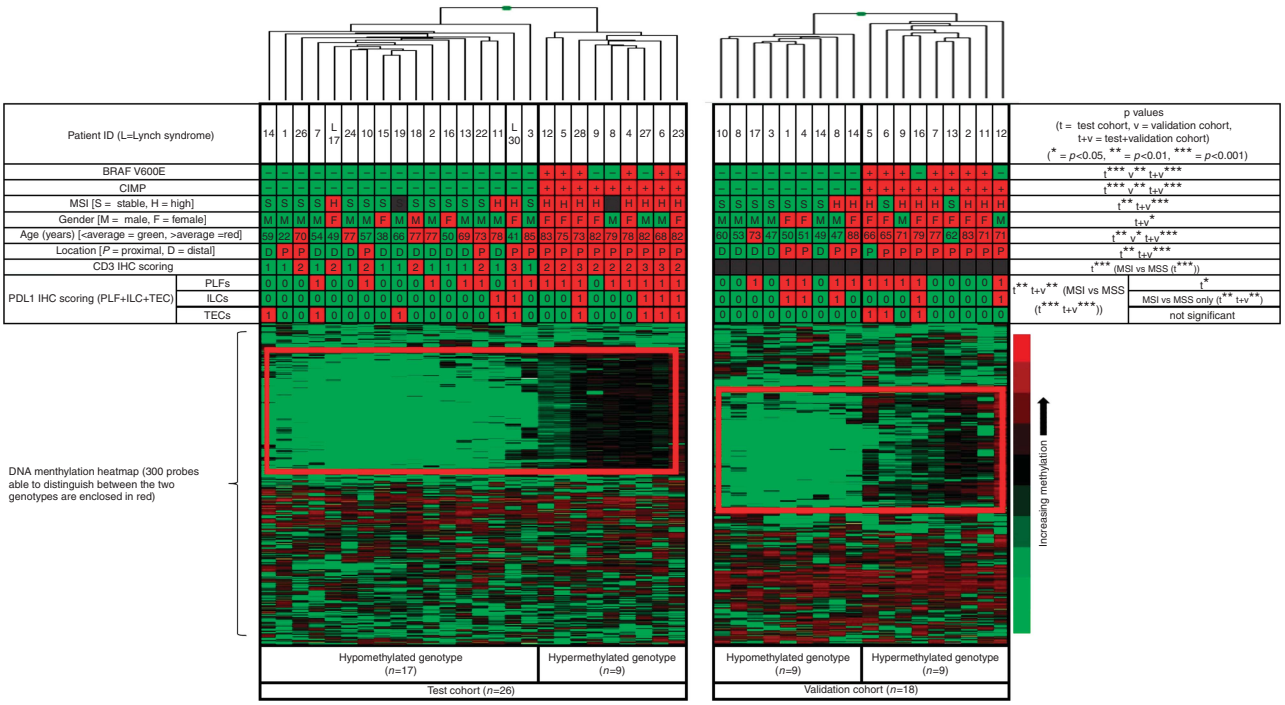

Figure 1. Alongside different DNA methylation patterns, data best able to distinguish between the hypomethylated and hypermethylated genotypes are shown in green and red respectively. ILC = intra-tumoural lymphoid cells; PLF = peritumoural lymphoid follicles; TEC = tumour epithelial cells.

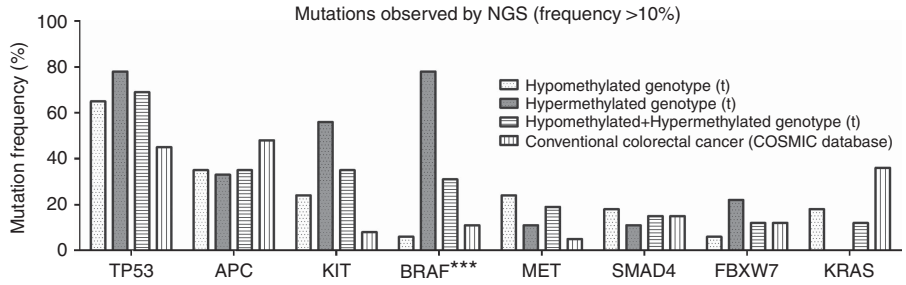

Figure 2. Mutations identified using NGS (>10% frequency). *BRAF* mutations are the most enriched in hypermethylated genotype and also overall compared to CRC average (t = test cohort, *** = $P<0.001$).

coefficient < -0.5 and average beta value difference >0.1 between normal and tumour samples).

Gene expression data was also used for CMS classification (Guinney *et al*, 2015). We observed CMS1 and CMS4 as the predominant subtypes in the hypermethylated (67%) and hypomethylated (53%) groups, respectively (Supplementary Table 7).

Microsatellite instability and PDL1 expression. MSI was called where three or more out of the five loci tested were observed to be aberrant. As shown in Figure 1, we observed most of the MSI cases in both the test and validation cohorts to fall within the hypermethylated group ($P<0.001$, $P=0.06$, $P<0.001$ in test, validation and both cohorts combined together respectively). Because it has recently been shown that metastatic MSI-high CRCs are good candidates for immune checkpoint inhibitor therapy, we tested for CD3 and PDL1 expression to evaluate the presence of adaptive immune resistance in our cohorts (Xiao and Freeman, 2015).

We initially conducted CD3 IHC to confirm the presence of an immune infiltrate in the test cohort. As shown in Supplementary Figure 3 we observed a higher infiltration of $CD3^+$ T-lymphocytes in both MSI cases compared to microsatellite stable cases (MSS) and also the hypermethylated group compared to hypomethylated group ($P<0.001$).

We then looked at PDL1 gene expression data and observed a higher expression of PDL1 in MSI cases compared to MSS (Figure 3A, $P=0.04$). We also observed a similar trend in the hypermethylated group compared to hypomethylated group; however it was not statistically significant (Figure 3A, $P=0.07$).

This finding was validated at the protein level using IHC in both the test and validation cohorts (representative staining can be seen in Figure 3B). A higher expression of PDL1 was observed in MSI cases compared to MSS cases ($P<0.001$, $P=0.16$, $P<0.001$ in test, validation and both cohorts combined together, respectively, Figure 3C). The trend was consistent across all the three cell populations we looked at but strongest in the ILCs ($P=0.003$, $P=0.3$, $P=0.001$ in test, validation and both cohorts combined together, respectively, Figure 3C). We also observed a similar trend comparing PDL1 protein expression in hypermethylated *vs* hypomethylated group ($P=0.008$, $P=0.8$, $P=0.03$ in test, validation and both cohorts combined together, respectively, Figure 3C).

Association between molecular and clinicopathological data. Patients in the hypermethylated group had a higher average age compared to the hypomethylated group ($P<0.01$ in test, validation and both cohorts combined together, Figure 1). We observed these to be mostly female patients ($P=0.10$, $P=0.15$, $P=0.01$ in test, validation and both cohorts combined together, respectively, Figure 1) and the tumours were mostly in the proximal colon ($P=0.01$, $P=0.13$, $P<0.01$ in test, validation and both cohorts combined together, respectively, Figure 1). We observed no statistically significant link between molecular data and any other clinicopathological parameters including prognosis (overall survival) even when adjusted for age/stage/gender/MSI/tumour location in a multivariate analysis (Supplementary Table 7).

DISCUSSION

Our study has for the first time identified two distinct genotypes within the SRCCa phenotype. Markers previously associated with this phenotype (e.g., *BRAF V600E* mutation, MSI and CIMP) are only present in one genotype, which represents only 41% of the cases in our cohorts (9/26 in test cohort, 9/18 in validation cohort). We also found this genotype to be associated with older age, female gender and predominant in the proximal colon. Again these are features which have been associated with the signet ring cell phenotype by a number of studies but our study shows that these only represent one (hypermethylated) genotype (Kakar *et al*, 2012; Barras, 2015). The study design is summarised in Supplementary Figure 4.

DNA methylation level was observed to be the major difference between the two genotypes with 202 genes (300 probes) splitting the cohorts into two groups. It is also interesting to see that all these genes follow the classic methylation pattern of CIMP genes, which are unmethylated in normal tissue. Methylation levels in the hypomethylated genotype are similar to those of the normal tissue and are only elevated in the hypermethylated genotype (data not shown). With the widespread availability of array-based methylation analysis it is now possible to look at methylation at a much deeper level than was possible when CIMP was discovered back in 1999 (Toyota *et al*, 1999). It can be seen from our data that the CIMP genotype in these cases comprises of a 202 gene signature as opposed to only five genes as has been traditionally thought. We also observed interesting differences in methylation patterns outside of the 300 probes that split the cohorts into two genotypes. The main difference was approximately twice the amount of methylation changes occurring within CpG islands in the hypermethylated genotype compared to hypomethylated genotype (44% *vs* 19%) (Supplementary Figure 5).

On the basis of NGS data we found *TP53*, *APC*, *KIT* and *BRAF* to be the most mutated genes. *TP53* and *APC* are also highly mutated in conventional CRC and *BRAF* is known to be frequently mutated in SRCCa. The finding of a common KIT mutation is novel. Across both the test and the validation cohorts (and across both the hypermethylated and hypomethylated genotypes) 25% of cases were found to carry a *KIT M541L* mutation. A Sanger sequencing assay confirmed this finding, and the high-incidence rate differs substantially from the minor allele frequencies reported in multiple databases

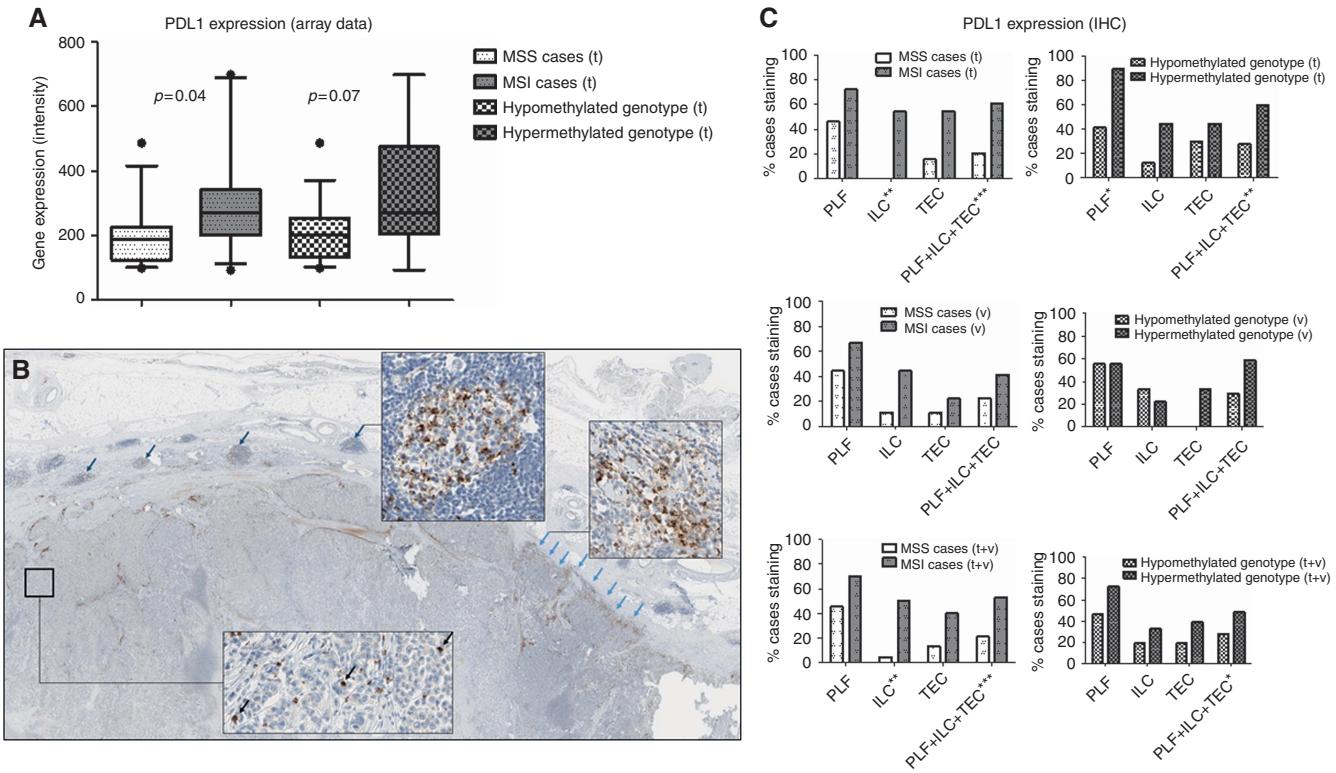

Figure 3. PDL1 expression. (**A**) PDL1 gene expression from array data. (**B**) Representative PDL1 staining at ×4 and ×40 in peritumoural lymphoid follicles (PLF) (dark blue arrows), intra-tumoural lymphoid cells (ILC) (black arrows) and tumour epithelial cells (TEC) along the invasive front (light blue arrows). (**C**) PDL1 IHC scoring in both cohorts ((*t*, *v* and *t*+*v*=test, validation and test+validation cohorts, respectively), (*, ** and *** = $P<0.05$, $P<0.01$ and $P<0.001$, respectively)).

(1000 Genomes frequency: 6.45% (The Genomes Project C, 2015), ExAC frequency: 7.89% (Lek *et al*, 2016), NHLBI ESP European frequency: 11.19% (National Heart, Lung, and Blood Institute). We know that *KIT* mutant gastrointestinal stromal tumours benefit from treatment with the tyrosine kinase inhibitor imatinib (Siehl and Thiel, 2007) and similarly it has been reported that this mutation not only increases proliferation but also enhances sensitivity to imatinib in certain cancers (Gonçalves *et al*, 2006; Masago *et al*, 2015; Iacono *et al*, 2016). Also it has not been reported previously in CRC and thus is of potential clinical significance as it may open new targeted approaches to treatment. This finding also highlights the distinct molecular profile of SRCCa and that it is not just an enrichment of conventional CRC.

Approximately 75% of cases in our study had stage III tumours, and we know that ~12% of all stage III colorectal tumours are MSI (Vilar and Gruber, 2010). However we observed 48% of our cases to be MSI, and most of them were in the hypermethylated genotype (88% of hypermethylated cases were MSI). We were also able to confirm the downregulation of MLH1 using gene expression data in this genotype, which indicates a defective DNA mismatch repair pathway (Kane *et al*, 1997) (Supplementary Figure 6). In light of recent developments highlighting the potential of immune checkpoint inhibitor therapies in MSI tumours, we also examined CD3 and PDL1 expression in our cohorts (Herbst *et al*, 2014; Xiao and Freeman, 2015). We observed higher CD3 and PDL1 levels in MSI cases compared to MSS (Figure 3). We also observed both these markers to be upregulated in the hypermethylated genotype compared to the hypomethylated genotype, suggesting that the hypermethylated genotype may potentially benefit from immune checkpoint inhibitor therapy because of the development of adaptive immune resistance (Figure 3; Llosa *et al*, 2015). This finding also fits in with recent studies, where upregulation of mTOR and MYC pathways (as observed in the hypermethylated genotype, Supplementary Figure 2 and Supplementary Table 5) can lead to PDL1 dependant suppression of the immune response (Casey *et al*, 2016; Lastwika *et al*, 2016). The immune checkpoint inhibitor therapy clinical trials in CRC have suffered from low-sample numbers because most MSI CRCs are early stage (Xiao and Freeman, 2015). This makes SRCCa hypermethylated genotype an ideal candidate for these trials as these cancers are likely to be both MSI and high stage (Le *et al*, 2015; Llosa *et al*, 2015).

Comparing our data to the CMS classification of Guinney *et al*, 2015 we find our hypermethylated group similar to the CMS1 (microsatellite instability immune, 14%) subtype with a high mutation count, MSI, CIMP, *BRAF* mutations, immune infiltration (as measured by CD3 IHC) and predominance in the proximal colon and the female gender (Guinney *et al*, 2015). The hypomethylated group shows similarities to CMS4 (mesenchymal, 23%) subtype in terms of upregulation of epithelial–mesenchymal transition genes, but also to CMS3 (metabolic, 13%) subtype as it contains all of the *KRAS* mutant tumours.

In summary, SRCCa comprises of two molecularly distinct genotypes. An MSI^+/$CIMP^+$/*BRAF V600E*$^+$/$CD3^+$/$PDL1^+$ hypermethylated genotype predominant in the proximal colon, and a hypomethylated genotype predominant in the distal colon. The high frequency of MSI and PDL1 expression in the hypermethylated genotype makes it a potential target for immune checkpoint inhibitor therapy. In addition, a high-detected frequency of the c.1621A>C (p.M541L) *KIT* actionable mutation also suggests imatinib as a candidate genomic targeted therapy. Testing tumour tissue for these two molecular aberrations may be

clinically beneficial upon making a diagnosis of SRCCa. Because of the rarity of this disease and the lack of cell line and animal models, the results of this study strongly support the need for an early phase trial aimed at these targets.

ACKNOWLEDGEMENTS

We would like to thank all patients whose samples were used in this study. We are also thankful to the Northern Ireland Biobank and Grampian Biorepository for providing us with tissue blocks and patient data; and Dr HG Coleman (Queen's University Belfast) for her advice on statistical analyses. This work has been carried out with financial support from Cancer Research UK (grant: C11512/A18067), Experimental Cancer Medicine Centre Network (grant: C36697/A15590 from Cancer Research UK and the NI Health and Social Care Research and Development Division), the Sean Crummey Memorial Fund and the Tom Simms Memorial Fund. The Northern Ireland Biobank is funded by HSC Research and Development Division of the Public Health Agency in Northern Ireland and Cancer Research UK through the Belfast CRUK Centre and the Northern Ireland Experimental Cancer Medicine Centre; additional support was received from Friends of the Cancer Centre. The Northern Ireland Molecular Pathology Laboratory which is responsible for creating resources for the Northern Ireland Biobank has received funding from Cancer Research UK, Friends of the Cancer Centre and Sean Crummey Foundation.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Alvi MA, McArt DG, Kelly P, Fuchs MA, Alderdice M, McCabe CM, Bingham V, McGready C, Tripathi S, Emmert-Streib F, Loughrey MB, McQuaid S, Maxwell P, Hamilton PW, Turkington R, James JA, Wilson RH, Salto-Tellez M (2015) Comprehensive molecular pathology analysis of small bowel adenocarcinoma reveals novel targets with potential for clinical utility. *Oncotarget* **6**(25): 20863–20874.

Anthony T, George R, Rodriguez-Bigas M, Petrelli NJ (1996) Primary signet-ring cell carcinoma of the colon and rectum. *Ann Surg Oncol* **3**(4): 344–348.

Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F (2017) Global patterns and trends in colorectal cancer incidence and mortality. *Gut* **66**(4): 683–691.

Barras D (2015) BRAF mutation in colorectal cancer: an update. *Biomark Cancer* **7**(Suppl 1): 9–12.

Bosman FT (2010) *WHO Classification of Tumours of the Digestive System*, 4 edn. IARC Press: Lyon, France.

Casey SC, Tong L, Li Y, Do R, Walz S, Fitzgerald KN, Gouw A, Baylot V, Guetegemann I, Eilers M, Felsher DW (2016) MYC regulates the antitumor immune response through CD47 and PD-L1. *Science* **352,**(Suppl 1): 227–231.

National Heart, Lung, and Blood Institute. *Exome Variant Server, GO Exome Sequencing Project (ESP)*. NHLBI: Seattle, WA, USA. Available at: http://evs.gs.washington.edu/EVS/ (Accessed April 2016).

Gonçalves A, Monges G, Yang Y, Palmerini F, Dubreuil P, Noguchi T, Jacquemier J, Di Stefano D, Delpero J-R, Sobol H, Bertucci F (2006) Response of a KIT-positive extra-abdominal fibromatosis to imatinib mesylate and KIT genetic analysis. *J Natl Cancer Inst* **98**(8): 562–563.

Gopalan V, Smith RA, Ho YH, Lam AK (2011) Signet-ring cell carcinoma of colorectum–current perspectives and molecular biology. *Int J Colorectal Dis* **26**(2): 127–133.

Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa EMF, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S (2015) The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**(11): 1350–1356.

Herbst RS, Soria J-C, Kowanetz M, Fine GD, Hamid O, Gordon MS, Sosman JA, McDermott DF, Powderly JD, Gettinger SN, Kohrt HEK, Horn L, Lawrence DP, Rost S, Leabman M, Xiao Y, Mokatrin A, Koeppen H, Hegde PS, Mellman I, Chen DS, Hodi FS (2014) Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**(7528): 563–567.

Iacono ML, Buttigliero C, Monica V, Bollito E, Garrou D, Cappia S, Rapa I, Vignani F, Bertaglia V, Fiori C, Papotti M, Volante M, Scagliotti GV, Porpiglia F, Tucci M (2016) Retrospective study testing next generation sequencing of selected cancer-associated genes in resected prostate cancer. *Oncotarget* **7**: 14394–14404.

Kakar S, Deng G, Smyrk TC, Cun L, Sahai V, Kim YS (2012) Loss of heterozygosity, aberrant methylation, BRAF mutation and KRAS mutation in colorectal signet ring cell carcinoma. *Mod Pathol* **25**(7): 1040–1047.

Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, Jessup JM, Kolodner R (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res* **57**(5): 808–811.

Lastwika KJ, Wilson 3rd W, Li QK, Norris J, Xu H, Ghazarian SR, Kitagawa H, Kawabata S, Taube JM, Yao S, Liu LN, Gills JJ, Dennis PA (2016) Control of PD-L1 Expression by Oncogenic Activation of the AKT-mTOR Pathway in Non-Small Cell Lung Cancer. *Cancer Res* **76**(2): 227–238.

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B, Diaz LAJ (2015) PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* **372**(26): 2509–2520.

Lek M, Karczewski K, Minikel E, Samocha K, Banks E, Fennell T, O'Donnell-Luria A, Ware J, Hill A, Cummings B, Tukiainen T, Birnbaum D, Kosmicki J, Duncan L, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Cooper D, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki M, Levy Moonshine A, Natarajan P, Orozco L, Peloso G, Poplin R, Rivas M, Ruano-Rubio V, Ruderfer D, Shakir K, Stenson P, Stevens C, Thomas B, Tiao G, Tusie-Luna M, Weisburd B, Won H-H, Yu D, Altshuler D, Ardissino D, Boehnke M, Danesh J, Roberto E, Florez J, Gabriel S, Getz G, Hultman C, Kathiresan S, Laakso M, McCarroll S, McCarthy M, McGovern D, McPherson R, Neale B, Palotie A, Purcell S, Saleheen D, Scharf J, Sklar P, Patrick S, Tuomilehto J, Watkins H, Wilson J, Daly M, MacArthur D (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**(7616): 285–291.

Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, Taube JM, Blosser RL, Fan H, Wang H, Luber BS, Zhang M, Papadopoulos N, Kinzler KW, Vogelstein B, Sears CL, Anders RA, Pardoll DM, Housseau F (2015) The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov* **5**(1): 43–51.

Masago K, Fujita S, Muraki M, Hata A, Okuda C, Otsuka K, Kaji R, Takeshita J, Kato R, Katakami N, Hirata Y (2015) Next-generation sequencing of tyrosine kinase inhibitor-resistant non-small-cell lung cancers in patients harboring epidermal growth factor-activating mutations. *BMC Cancer* **15**(1): 1–8.

McCourt CM, McArt DG, Mills K, Catherwood MA, Maxwell P, Waugh DJ, Hamilton P, O'Sullivan JM, Salto-Tellez M (2013) Validation of next generation sequencing technologies in comparison to current diagnostic gold standards for BRAF, EGFR and KRAS mutational analysis. *PLoS ONE* **8**(7): e69604.

Nitsche U, Zimmermann A, Späth C, Müller T, Maak M, Schuster T, Slotta-Huspenina J, Käser SA, Michalski CW, Janssen K-P, Friess H, Rosenberg R, Bader FG (2013) Mucinous and signet-ring cell colorectal cancers differ from classical adenocarcinomas in tumor biology and prognosis. *Ann Surg* **258**(5): 775–782.

Siehl J, Thiel E (2007) C-kit, GIST, and imatinib. *Recent Results Cancer Res* **176**: 145–151.

The Genomes Project C (2015) A global reference for human genetic variation. *Nature* **526**(7571): 68–74.

Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP (1999) CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci USA* **96**(15): 8681–8686.

Vilar E, Gruber SB (2010) Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol* **7**(3): 153–162.

Xiao Y, Freeman GJ (2015) The Microsatellite Instable (MSI) subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discov* **5**(1): 16–18.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)