

An immunogenomic stratification of colorectal cancer: Implications for development of targeted immunotherapy

Neeraj Lal, Andrew D Beggs, Benjamin E Willcox, and Gary W Middleton*

Cancer Immunology and Immunotherapy Centre; School of Cancer Sciences; University of Birmingham; Birmingham, UK

Keywords: colorectal cancer, immune signature, microsatellite instability, *RAS*, stratification

Abbreviations: CIRC, Co-ordinate Immune Response Cluster; CMS, consensus molecular subtypes; CRC, colorectal cancer; CRCSC, colorectal cancer subtyping consortium; MSI-H, microsatellite unstable (high); MSI-L, microsatellite unstable (low); MSS, microsatellite stable; TCGA, The Cancer Genome Atlas; TILs, tumor-infiltrating lymphocytes

Although tumor infiltrating lymphocyte (TIL) density is prognostic and predictive in colorectal cancer (CRC), the impact of tumor genetics upon colorectal immunobiology is unclear. Identification of genetic factors that influence the tumor immunophenotype is essential to improve the effectiveness of stratified immunotherapy approaches. We carried out a bioinformatics analysis of CRC data in The Cancer Genome Atlas (TCGA) involving two-dimensional hierarchical clustering to define an immune signature that we used to characterize the immune response across key patient groups. An immune signature termed The Co-ordinate Immune Response Cluster (CIRC) comprising 28 genes was coordinately regulated across the patient population. Four patient groups were delineated on the basis of cluster expression. Group A, which was heavily enriched for patients with microsatellite instability (MSI-H) and POL mutations, exhibited high CIRC expression, including the presence of several inhibitory molecules: *CTLA4*, *PDL1*, *PDL2*, *LAG3*, and *TIM3*. In contrast, *RAS* mutation was enriched in patient groups with lower CIRC expression. This work links the genetics and immunobiology of colorectal tumorigenesis, with implications for the development of stratified immunotherapeutic approaches. Microsatellite instability and POL mutations are linked with high mutational burden and high immune infiltration, but the coordinate expression of inhibitory pathways observed suggests combination checkpoint blockade therapy may be required to improve efficacy. In contrast, *RAS* mutant tumors predict for a relatively poor immune infiltration and low inhibitory molecule expression. In this setting, checkpoint blockade may be less efficacious, highlighting a requirement for novel strategies in this patient group.

Introduction

The density of TILs and the expression of certain immune-related genes are of prognostic and predictive value in CRC.^{1–5} However, the factors that determine a patient's immune phenotype are unclear, and few systematic analyses have investigated the somatic and germline molecular drivers of immune infiltration. Although, microsatellite unstable (MSI-H) cancers are known to be associated with increased TIL density,^{6–12} the nature of the immune infiltration and the molecular drivers of the immune phenotype in microsatellite stable (MSS) CRC are poorly understood. In particular, it is unclear whether defined molecular subsets (*RAS* mutant, *BRAF* mutant, *PIK3CA* mutant, quadruple wildtype (*BRAF*, *PIK3CA*, *NRAS*, *KRAS* all wildtype)) are associated with high or low immune infiltration. In addition,

in both MSI-H and MSS cancers, the extent to which therapeutically tractable inhibitory immune checkpoint receptors are represented is unclear and is of substantial interest, particularly considering recent checkpoint blockade failures in CRC.^{13,14} Given the prognostic and predictive relevance of CRC immunophenotype, a clearer understanding of the link between immunophenotype with tumor genotype is crucial.

In this study, we carried out a bioinformatics analysis of CRC data in The Cancer Genome Project (TCGA). Our analyses defined a co-regulated cluster of immune related genes with a distinct *Th1* bias, expression of which defines four patient subgroups. The findings have implications for our understanding of the immunobiology of colorectal carcinogenesis, and for development of stratified immunotherapy approaches for CRC.

© Neeraj Lal, Andrew D Beggs, Benjamin E Willcox, and Gary W Middleton

*Correspondence to: Gary W Middleton; Email: g.middleton@bham.ac.uk

Submitted: 09/04/2014; Revised: 10/08/2014; Accepted: 10/09/2014

<http://dx.doi.org/10.4161/2162402X.2014.976052>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

Results

Coordinate expression of immune response-related genes in colorectal cancer

We interrogated expression of immune response-related genes in the CRC dataset in the TCGA ($n = 195$). Preliminary analyses focused on an initial gene list (Table S1) based partly on previous studies on bio-molecular networks incorporating immune genes linked with disease-free survival in CRC.^{1,4,5} These included those associated with *Th1* subset function (*STAT1*, *IRF1*, *IFNG*, *TBX21*, *IL18RAP*, *ICOS*, *GNLY*), certain chemokines (*CX3CL1*, *CXCL9*, *CXCL10*), adhesion molecules (ICAM and MADCAM), and an array of class II genes. Also, we included a number of immune checkpoint genes (PD-1, PD-L1, PD-L2, LAG3, TIM3, CTLA-4), two of which (PD-L1 and PD-L2) were previously associated with outcome.⁴ Finally, our initial gene list was supplemented with class I genes and additional class II genes and genes involved in T cell activation, together with NKG2D ligands and related genes relevant to innate immune activation (e.g. ULBPs, *PROCR*¹⁵).

We performed unsupervised two-dimensional hierarchical clustering to assess the extent to which gene expression was coordinate or independent across the patient cohort. Visual analysis of the clustering highlighted a 28-gene subset (Table 1) which formed a clear gene grouping, expression of which was coordinately regulated (Fig. S1). Gene-tree analysis of the dendrogram confirmed the validity of this grouping, identifying a subset of 24 highly coordinated genes (distance threshold 0.46)

Table 1. Genes within the coordinate immune response cluster (CIRC). These genes are presented in the order of the CIRC signature

Gene ID
<i>HLA-DQA1</i>
<i>HLA-DQA2</i>
<i>HLA-DRB5</i>
<i>CTLA4</i>
<i>PDCD1LG2</i>
<i>ICAM1</i>
<i>CD274</i>
<i>STAT1</i>
<i>IRF1</i>
<i>IFNG</i>
<i>GNLY</i>
<i>TBX21</i>
<i>CCL5</i>
<i>LAG3</i>
<i>CD247</i>
<i>ICOS</i>
<i>IL18RAP</i>
<i>CXCL9</i>
<i>CXCL10</i>
<i>HLA-DPB1</i>
<i>HLA-DPA1</i>
<i>HLA-DMB</i>
<i>HLA-DRA</i>
<i>HLA-DMA</i>
<i>CD80</i>
<i>HLA-DOA</i>
<i>CD4</i>
<i>HAVCR2</i>

predominantly associated with *Th1* immunity, including numerous class II MHC loci, and inhibitory molecules targeted in checkpoint blockade strategies. This 24-gene block was identical to our 28-gene cluster other than the absence of three additional class II MHC loci (*HLA-DRB5*, *HLA-DQA2*, and *HLA-DQA1*) and one additional inhibitory molecule (*HAVCR2* (TIM3)), all of which correlated closely with the 24-gene block and were positioned directly adjacent on the gene cluster dendrogram. We therefore, proceeded with the 28-gene cluster as it formed a clearly coordinated block on visual and correlation analysis. We termed this grouping the CIRC.

Of the 28 genes in the CIRC, 20 have previously been associated with outcome based on experimental data.^{1,4,5} We show here that these prognostic immune response genes are highly coordinately expressed in CRC and correlate with other biologically associated genes including *HLA-DQA1*, *HLA-DRB5*, *HLA-DPB1*, *LAG3*, *TIM3*, *CTLA4*, and *CCL5*. The degree of correlation between immune checkpoint receptor/ligand gene expression was notable (Table S2). Particularly striking was the correlation between PD1 and LAG3 ($r^2 = 0.62$, Fig. 1). Given the Th1 bias of the cluster we performed separate unsupervised hierarchical clustering analysis of Th2 cytokines IL2, IL3, IL4, IL5, and IL6 and transcription factor GATA3 (data not shown), revealing that these genes were excluded from the CIRC. We also note that the Th17 cytokine IL17A was also excluded from the cluster (Fig. S1).

Molecular determinants of the co-ordinate immune response cluster

Two-dimensional hierarchical clustering was performed for the entire gene list (Table S1) incorporating expression data together with molecular and clinical characteristics (*KRAS*, *BRAF*, *NRAS*, *TP53*, *PIK3CA*, and *PTEN* mutations; microsatellite status, methylation subtype, tumor stage, tumor site, and recurrence data). The microsatellite status of the total patient population was as follows: 11.8% were MSI-H, 15.9% were MSI-L (microsatellite-low), and 71.8% were MSS. 6.7% of patients were *BRAF* mutant (*MT*)/*MSI-H*, and 3.1% were *BRAF MT/MSS*. The mutational status of the cohort was as follows – *KRAS MT* (39.5%), *NRAS MT* (8.2%), *BRAF MT* (9.7%), *PIK3CA MT* (18.4%), and *TP53 MT* (51.8%). 38.5% of patients were quadruple wild type (*BRAF WT*, *KRAS WT*, *NRAS WT* and *PIK3CA WT*). We also investigated mutations in the novel genes *POLE* (7.2% of patients) and *POLD1* (2.6% of patients), which have recently been highlighted as key drivers of colorectal carcinogenesis for a minority of CRC patients.¹⁶

Hierarchical clustering delineated four distinct patient groups (Fig. 2, Table 2). Group A patients demonstrated strong CIRC expression (mean expression 0.98). Notably, all of the MSI-H patients were included in this group, who constituted 82.1% of all group A patients, whereas only 10.7% were MSS and 7.1% were MSI-L. Across the entire cohort, expression of the CIRC signature was significantly higher in the MSI-H cancers vs. MSS ($p < 0.001$) and MSI-L cancers ($p < 0.001$). Multivariate analysis revealed that expression of *HLAA*, *HLAB*, and *HLAC* were all significantly less in MSI-H than MSS cancers ($p = 0.027$, $p =$

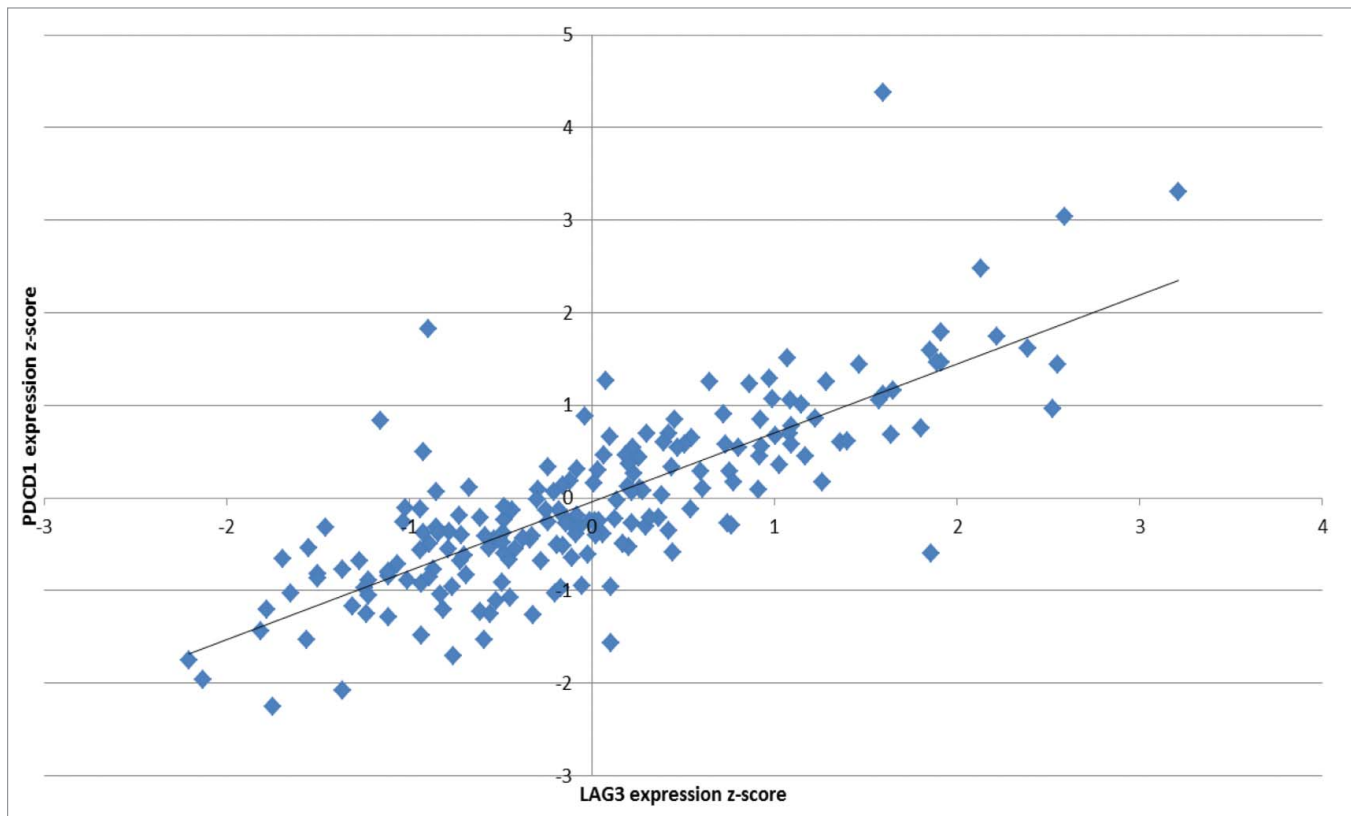


Figure 1. Significant correlation between *LAG3* and *PDCD1* (*PD1*) expression. Pearson correlation analysis indicates the inhibitory molecules *LAG3* and *PDCD1* (*PD1*) are highly coordinated in mRNA expression ($R^2 = 0.623$).

0.017, and $p = 0.018$, respectively for *HLA-A*, *B* and *C*), consistent with previous observations.¹⁷ Group A was characterized by the CpG island methylator phenotype (CIMP-H) (67.9%), right-sided tumor site (82.1%), *TP53* wild type (65%) and *KRAS* wild type (82.1%). 50% of patients in this group were *BRAF* mutant and of these *BRAF* mutants, 92.9% were MSI-H. Notably, *POLE* and *POLD1* mutant tumors were associated with higher expression of the CIRC ($p < 0.05$). 42.9% of patients in group A were either *POLE* or *POLD1* MT, and 70.6% of all patients with *POLE* or *POLD1* mutations were found in group A; 100% of *POLD1* mutants were assigned to group A. Of the non-MSI-H patients in group A, 3/5 were *POLE* mutant, two of which were MSS and one MSI-L.

We then analyzed patient groups B, C, and D (Fig. 2), which displayed lower expression of the CIRC signature than group A, to determine whether any molecular characteristics were associated with low CIRC expression. Group D comprised cancers with the lowest expression of the cluster (mean expression -0.62) and represented 43% of the entire patient population; group B had a lower mean cluster expression (0.13) than group C (0.50). Significantly, *RAS* mutation (*KRAS/NRAS*), occurring in 47.7% of patients, was associated with lower CIRC expression ($p < 0.001$). This was confirmed in the provisional TCGA data set, in which mutation and expression data for a further 30

patients (225 total) is available ($p < 0.001$). The frequency of *RAS* mutation was significantly higher in group B and D than other patient groups ($p = 0.01$), with group D having the highest frequency (D (61.9%) > B (51.0%) > C (28.1%) > A (21.4%)), ($p = 0.01$). Consistent with these observations, within MSS cancers (enriched in groups B–D), there was a strong trend for *RAS* mutation to be associated with low CIRC expression ($p = 0.076$). In group A, 83.3% of *RAS* mutations were MSI-H and 16.7% were MSS (Table 2). In *RAS* mutant MSI-H patients, the expression of *HLA-DRA* was lower (0.518) than in *RAS* wild types (1.027) but this difference did not reach significance ($p = 0.209$). Finally, multivariate analysis indicated that *KRAS* and *NRAS* mutation was associated with decreased *CD4* expression ($p < 0.001$) (Table S3). The association of low expression of the cluster and *RAS* mutation was particularly strong in the case of *NRAS* mutation where 76.9% of these cancers were in group D and over 90% of these tumors clustered in the two lowest expression groups (B and D). Thus *RAS* mutation, unless linked with microsatellite instability, is commonly associated with low levels of *Th1* infiltration and activation, class II expression and inhibitory checkpoint expression. Table 3 shows the distribution of key mutations.

Several groups have identified molecular subtypes of CRC based on expression data. In an effort to create consensus, The

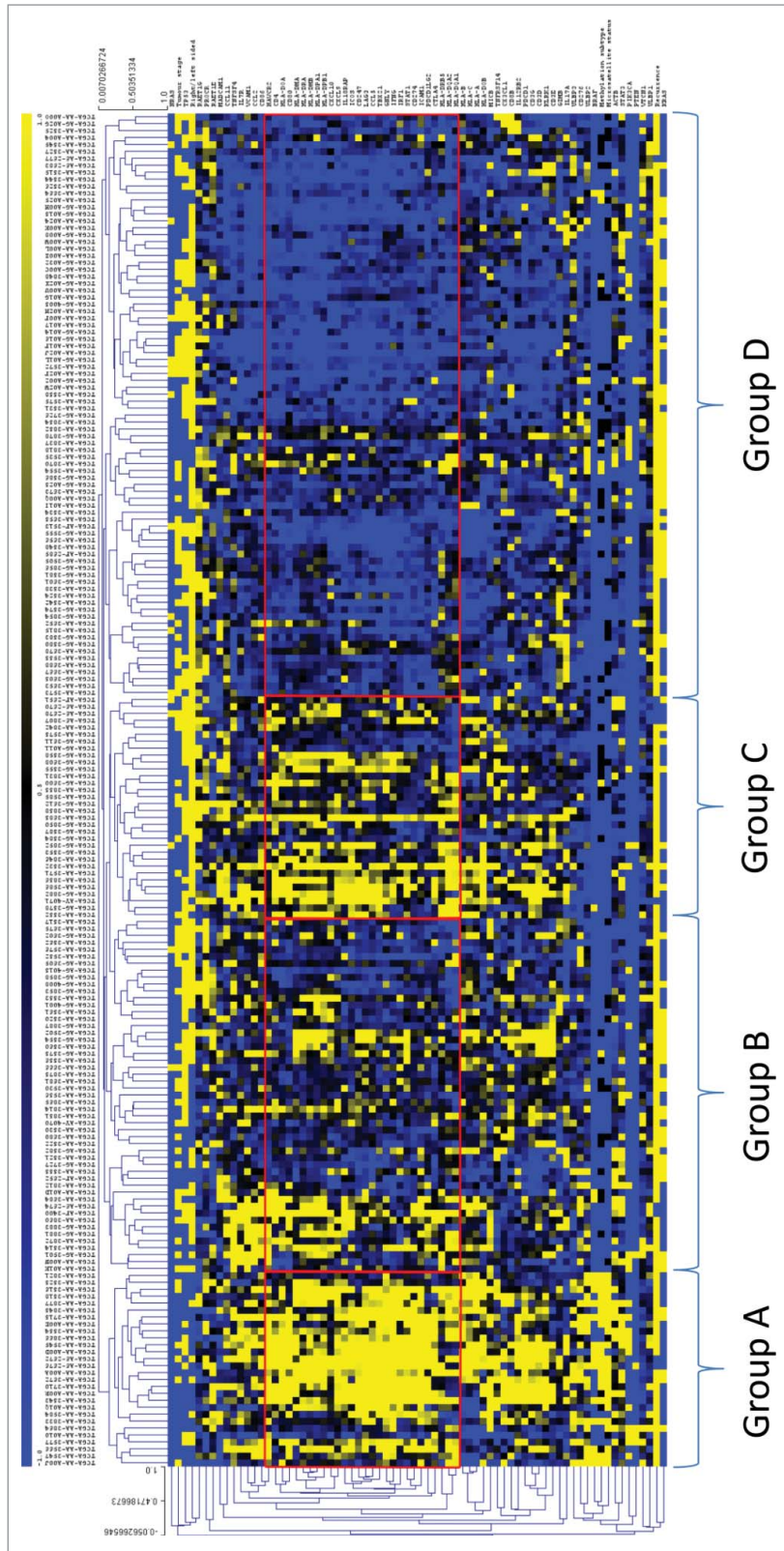


Figure 2. Two-dimensional hierarchical clustering delineates distinct immunological CRC patient groups. Gene expression (yellow, high expression; black, intermediate; blue, low expression) was clustered together with mutation data of key genes (*TP53*, *KRAS*, *BRAF*, *NRAS*, *PI3KCA* and *PTEN* (yellow, mutant; blue, wildtype)) and clinical data (microsatellite status (yellow, MSI-H; black, MSI-L; blue, MSS), recurrence data (yellow, recurred/progressed; blue, disease-free), tumor site (yellow, left sided; blue, right sided), tumor stage (yellow, stage III/IV; blue, stage I/II), methylation subtype (yellow, CIMP-H; black, CIMP-L; blue, CIMP-negative)). Clustering was performed by genes/mutations/clinical data (rows) and patients (columns) using the Pearson algorithm. Red boxes indicate groups of patients with strong clustering of the coordinate immune response cluster. Patients were delineated into four distinct groups (A–D) on the basis of the dendrogram and the cluster expression.

Table 2. Characteristics of patient groups. For microsatellite status, methylation and tumor side, the most frequent result is stated. Tumor side refers to the right or left side of the colon. The cluster expression pattern displays the expression pattern of the CIRC cluster in each patient group

	Group A	Group B	Group C	Group D
Microsatellite status	MSI-H (82%)	MSS (86%)	MSS (94%)	MSS (75%) MSI-L (25%)
Methylation	CIMP-High (68%)	CIMP-Neg (77%)	CIMP-Neg (66%)	CIMP-Neg (69%)
Side of tumour	Right (82%)	Left (63%)	Left (94%)	Left (82%)
Stage I+II	73%	55%	72%	49%
TP53 MT	35%	65%	62%	48%
BRAF MT	50%	4%	3%	1%
KRAS MT	18%	47%	22%	49%
NRAS MT	0%	4%	9%	13%
PIK3CA MT	39%	14%	9%	18%
Quadruple WT	18%	39%	69%	33%
Percentage of patients	14%	26%	16%	43%
Mean cluster expression z-score	+0.98	+0.13	+0.50	-0.62
Cluster expression pattern				

colorectal cancer subtyping consortium (CRCSC) have defined an integrated classification – consensus molecular subtypes (CMS).¹⁸ To investigate how closely our CIRC patient groupings (based predominantly on immunological gene expression) correlate with the recently announced CMS classifications, we retrieved CMS data for the TCGA dataset. **Table 4** shows the distribution of CMS classifications across our four patient groups. As expected, CMS1, which is a group comprised of mostly MSI-H and *BRAF MT* tumors with strong immunity, falls predominantly within Group A. Interestingly, 80% of CMS3 tumors, which are epithelial with *KRAS* mutations, fall within patient groups B and D, similar to our *KRAS* distribution data (**Table 3**). This reinforces the hypothesis that *KRAS* mutant tumors are immunologically bland. A large proportion of CMS4 tumors, which are

predominantly mesenchymal with a poor prognosis, were found in patient group B (63.9%), and a large proportion of CMS2 tumors, which are epithelial with high CIN and WNT/MYC pathway activation, fall within patient group D (61.5%). This demonstrates that patient groups B and D are distinct despite both groups having low CIRC expression levels.

Discussion

In this study, we identified a group of tightly co-regulated immune-related genes that we termed the CIRC, and used this to assess differences in the intra-tumoral immune response in a molecularly characterized cohort of CRC patients. The CIRC

Table 3. Frequencies of mutations in total patient population and distribution of mutations across patient groups A–D. We classified patients initially on the basis of *KRAS* status, followed by *NRAS*, *BRAF*, and finally *PIK3CA*. According to these criteria, *KRAS* mutations may also have *NRAS*, *BRAF*, and *PIK3CA* mutations (*KRAS* MT +/-*NRAS* MT +/-*BRAF* MT +/-*PIK3CA* MT). *NRAS* mutations are *KRAS* wildtype but may have mutations in *BRAF* and *PIK3CA* (*KRAS* WT + *NRAS* MT +/-*BRAF* MT +/-*PIK3CA* MT). *BRAF* mutants are *KRAS* and *NRAS* wildtype but may have mutations in *PIK3CA* (*KRAS* WT + *NRAS* WT + *BRAF* MT +/-*PIK3CA* MT). *PIK3CA* mutants are wildtype for *KRAS*, *NRAS*, and *BRAF*. Quadruple wildtype patients are *KRAS*, *NRAS*, *BRAF*, and *PIK3CA* wildtype. *TP53* mutation status is independent of other mutations

	% of total patient population	% of total in patient group A	% of total in patient group B	% of total in patient group C	% of total in patient group D
<i>TP53</i> MT	51.8%	6.9%	32.7%	19.8%	40.6%
<i>BRAF</i> MT	9.2%	72.2%	11.1%	5.5%	11.1%
<i>KRAS</i> MT	41.0%	7.5%	30%	10%	52.5%
<i>NRAS</i> MT	6.7%	0%	15.4%	7.7%	76.9%
<i>KRAS/NRAS</i> MT	47.7%	6.5%	28.0%	9.7%	55.9%
<i>PIK3CA</i> MT	4.1%	50%	37.5%	0%	12.5%
Quadruple wildtype	39.0%	6.6%	26.3%	28.9%	38.2%

Table 4. Distribution of Consensus Molecular Subtype (CMS) groups across CIRC patient groups. The percentage of each CMS group that fall within each CIRC patient group. 5.1% of TCGA patients did not have CMS classification data available

	Percentage of total patient population	CIRC Group A	CIRC Group B	CIRC Group C	CIRC Group D
CMS1	13.8%	81.5%	7.4%	3.7%	7.4%
CMS2	40.0%	0.0%	15.4%	23.1%	61.5%
CMS3	10.3%	0.0%	30.0%	20.0%	50.0%
CMS4	18.5%	5.6%	63.9%	13.9%	16.7%
Unclassified	12.3%	12.5%	25.0%	12.5%	50.0%

signature we defined stems from the work of Galon and colleagues, who first established the prognostic impact of T cell infiltration in CRC, initially highlighting seven co-modulated genes principally associated with *Th1*-associated immunity that correlated with outcome.¹ Subsequently, they demonstrated that the *Th1* genes, *TBX21*, *IFNG*, *IRF1* and *STAT1* were all individually associated with outcome.⁴ Other independent immune gene predictors in that analysis were *IL18RAP*, *ICOS*, PD-L1, PD-L2 and PD-1. We extended this gene list by adding the class I and class II genes, further immune checkpoint genes and a number of other genes related to innate immune recognition and T cell activation, generating an initial gene list of > 50 genes. Unsupervised hierarchical clustering then highlighted a core cluster of 28 tightly co-regulated genes comprising the CIRC signature. Importantly, 20 of the 28 genes had previously been highlighted by Galon and colleagues^{1,4,5} as either prognostically or predictively relevant, including five of the original prognostic seven gene cluster (with only CD8A (for which microarray data was unavailable) and GZMB not represented). In contrast, genes associated with Th2 and Th17 profiles were excluded from the cluster. Thus, applying a different approach to an independent data-set our analysis verified the high level of correlation of a number of genes based on *Th1*-associated immunity, and also highlighted involvement of additional immune genes.

An important feature of the CIRC signature is that it includes essentially all class II MHC loci, as well as CD4, whereas in contrast, expression of class I MHC molecules, CD8B and also GZMB are all excluded from the signature. In addition to the critical role of T helper cells for CD8⁺ cell priming^{19,20} and expansion,^{21, 22} CD4⁺ cells have also been suggested to be major mediators of immunological tumor cell death.²³⁻²⁵ Adoptive CD4⁺ T cell transfer has been found to upregulate class II expression on tumor cells mediating protection from tumor progression.²⁴ Class II upregulation was mediated via IFN γ and protection was attenuated using anti-IFN γ antibodies. Recent adoptive transfer approaches involving autologous CD4⁺ T cells have met with clinical success. A durable complete response was obtained in a patient with melanoma after infusion of NY-ESO-1 specific CD4⁺ cells recognizing an HLA-DP4 restricted epitope.²⁶ Additionally, a durable response was obtained in a patient with cholangiocarcinoma on infusion of autologous mutation-specific CD4⁺ cells which adopted a poly-functional *Th1* phenotype.²⁷ These studies highlight the potential for CD4⁺ cells to mediate clinically potent anti-tumor responses via *Th1* mechanisms. The CD4-centric nature of our CIRC highlights that CD4⁺ T cells may be important in CRC anti-tumor immunity.

The CIRC included the major immune checkpoint molecules. Not only PD-L1, PD-L2 but also LAG3, TIM3 and CTLA4 were all represented in the CIRC and there was a high degree of correlation between inhibitory checkpoint gene expression. This is consistent with the expected feedback sequelae of a pronounced *Th1* infiltrate. IFN γ is the canonical cytokine associated with Th1 T helper cells and expression of PD-L1 is significantly augmented by IFN γ .²⁸ *IRF1* is of primary importance in the constitutive expression of PD-L1 and in IFN γ -driven upregulation.²⁹ In animal models, significant percentages of infiltrating CD4⁺ and CD8⁺ cells co-express high levels of both PD-1 and *LAG-3*.³⁰ Dual anti-LAG3/anti-PD-1 immunotherapy strikingly enhanced survival in the MC38 colon cancer model compared with animals treated with single antibodies alone.³⁰ Our data provide justification for trialing combinations of checkpoint blockade agents in CRC.

One of the key aims of this study was to investigate the somatic factors associated with the immune response in CRC. Group A, which exhibited strong expression of the CIRC signature, was dominated by MSI-H tumors, all of which fell in this grouping. This complements very well the work of Bissière-Michot and colleagues, who revealed that MSI-H tumors have a high density of *Tbet*-positive Th1 T cells relative to MSS tumors, and higher expression of the chemokines *CCL5*, *CXCL9*, and *CXCL10*, all of which are found in the CIRC.¹² The *Th1* response may be driven through activation of the *CXCL9/CXCL10* signaling axis.¹² Previous studies have also shown that MSI-H phenotype is highly associated with class II expression.³¹ Strongly DR positive tumors had a significantly higher TIL density than those with absent or weak staining and survival was significantly better in patients with high DR expression. Consistent with this, Bindea and colleagues demonstrated that the expression of several class II genes, including *HLA-DRA*, are individually associated with improved disease-free survival.⁵ High DR expression in MSI-H CRC contrasts with that of class I molecules, which are completely lost in 60% of sporadic MSI-H cases and only 16.7% of right sided MSS cancers.¹⁷ Thus, the immune landscape of MSI-H CRC, which is characterized by a high mutational burden including frameshift mutations,¹⁰ is dominated by T helper cell infiltration and activation, class II expression and coordinated upregulation of a range of immune checkpoint genes. These patients may therefore have greater responses to checkpoint blockade therapies. Similarly, *POLE/POLD1* mutant tumors, which also have a high mutational burden,¹⁶ were also associated with high CIRC expression. These data support the hypothesis that high

mutational burden may translate into a strong immune response, possibly due to neoantigen presentation, leading to better outcomes. The finding that the gene encoding TGF β R undergoes frameshift mutation in 90% of MSI-H CRC, giving rise to a highly immunogenic promiscuous class II peptide, is entirely in keeping with this hypothesis.³²

RAS mutation was significantly associated with lower CIRC expression, with over 60% of cancers in the group exhibiting very low expression (group D) being *RAS* mutant. *KRAS* and *NRAS* mutant CRC had significantly lower levels of CD4⁺ T cells on multivariate analysis. Although 21.4% of cancers in group A were *RAS* mutant, over 80% of these were MSI-H. This is the first analysis to our knowledge to clearly define the immunological landscape of *RAS* mutant CRC. Ogino and colleagues examined the interaction between T cell infiltration and *KRAS*, *BRAF*, and *PIK3CA* mutation status and found no significant associations with *KRAS* mutation,³³ but this study was limited to the density of CD3⁺, CD8⁺, CD45RO⁺, and *FoxP3*⁺ cells. Morris and colleagues demonstrated that transfection of mutant *KRAS* into fibroblasts abrogated the IFN γ -mediated upregulation of class II expression: effects on class I expression were minimal.³⁴ *RAS* transfection inhibited proliferation and IFN γ production of alloreactive T cells, an effect mediated by loss of class II expression on target cells.³⁵ These data suggest a possible mechanism for our observed paucity of class II expression and hence *Th1*-related gene expression in *RAS* mutant CRC. Our data complement previous studies demonstrating widespread abnormalities of class I presentation in *RAS* mutant CRC.³⁶ In MSI-H tumors, it is likely that any inhibitory effect of *RAS* mutation is overcome by the strong immunity resulting from neoantigens linked to mutations induced by microsatellite instability. Thus MSI-H *RAS* mutant tumors retain strong *Th1* immunity, despite having numerically lower *HLA-DRA* expression than MSI-H *RAS* wildtype tumors. In sum, these data suggest that the micro-environment of *RAS* mutant CRC is relatively immunologically unfavorable to conventional $\alpha\beta$ T cell responses.

An important consideration is how the CIRC signature relates to immunosuppressive pathways. Though *FOXP3* was not in the hierarchical clustering due to an absence of microarray data, we analyzed RNA sequencing data which showed that *FOXP3* expression is highest in patient group A and lowest in group D (data not shown). Additionally, the highly immunosuppressive cytokine IL10 also exhibited high expression in this patient group (data not shown), both consistent with increased inhibitory immunoregulation following establishment of a *Th1*-polarized microenvironment. We also analyzed expression of *STAT3*, which has been linked to enhanced levels of tumor-associated macrophages, myeloid-derived suppressor cells, and poor T cell responses.^{37,38} Although not strongly linked to the CIRC, we found increased *STAT3* expression was associated with *PIK3CA* mutation (Table S3). Conditioned media from *PIK3CA* mutant cells has been found to drive *STAT3* activation³⁹; Aspirin decreases *STAT3* activation in murine CRC models and promotes apoptosis. Our findings may therefore partly explain previous molecular epidemiology studies⁴⁰⁻⁴², which have demonstrated that aspirin usage reduces the risk of CRC

recurrence following resection in patients with CRC harboring *PIK3CA* mutations.

The CIRC molecular associations relate to the CRCSC CMS and their immune and molecular characteristics.¹⁸ The majority of CMS1 patients (MSI high and immune activated) fall within CIRC group A. 82.5% of *KRAS* mutated patients were in CIRC groups D and B (the lowest two CIRC expression groups) and almost the same percentage of CMS3 patients, characterized by *KRAS* mutations and no significant immune infiltration and activation, were present in CIRC groups D and B.

In conclusion, we have shown that MSI-H and POL mutant CRC is associated with high-level expression of a coordinated immune response cluster characterized by T helper cell and class II related genes together alongside a range of chemokines and immune inhibitory checkpoint molecules. In contrast, *RAS* mutant tumors have significantly lower expression of this cluster. These findings have potential therapeutic implications. MSI-H tumors may be particularly amenable to CD4⁺ cell expansion and adoptive transfer approaches. Second, although there is interest in the use of immune checkpoint inhibitors in CRC, the coordinated upregulation of immune checkpoint genes suggests that a combinatorial approach is likely to be more successful. Finally, any immunologically based therapy in MSS CRC aimed at *RAS* mutant patients must take into account the relatively immunologically quiescent status of the micro-environment in the majority of *RAS* mutant CRC.

Methods

Gene identification

Genes associated with innate and adaptive tumor immunity in cancer^{1,4} and genes relevant to established colon cancer pathways were shortlisted as initial genes. This list was expanded using molecular network pathway analysis.⁴³⁻⁴⁵ This revealed further molecules and genes with known genetic, pathway and functional associations with our initial genes. Genes revealed through this approach were screened for inclusion into our final gene list based on known immunological relevance (Table S1).

Data extraction

Normalized Agilent microarray and RNAseq z-score data for initial genes and mutation data for *TP53*, *KRAS*, *BRAF*, *NRAS*, *HRAS*, *PIK3CA* and *PTEN* were extracted from TCGA colorectal data set⁴⁶ using the cBioportal tool.⁴³⁻⁴⁵ Where expression data were unavailable, genes were excluded from the analysis. Clinical data for these patients were retrieved from the TCGA portal (<https://tcga-data.nci.nih.gov/tcga/>) and tabulated with genetic data. CMS data were obtained from the CRCSC through the Synapse platform (<https://www.synapse.org/>, Sage Bionetworks).

Data analysis

Data were tabulated in Excel (Microsoft Corp.) and unsupervised two-dimensional hierarchical clustering was performed using MeV (Dana-Farber Cancer Institute, Boston, MA, USA) and the Pearson correlation. Patient and gene clusters were

identified by varying gene-tree distance-thresholds and visual analysis. Normality of distributions was confirmed with the Anderson–Darling test. Pearson Coefficient of determination (R^2) values were calculated in Excel and Minitab (Minitab Inc.) to investigate correlations in gene expression. Gene expression was correlated against β actin as a control gene. T tests were performed for univariate analyses of normally distributed data, and Mann–Whitney U tests for non-normally distributed data. Multivariate linear regression analyses of gene expression against mutations in key tumor associated genes (*P53*, *KRAS*, *BRAF*, *NRAS*, *HRAS*, *PI3KCA*, *PTEN*) were completed in Stata 12.2 (Statacorp.). For two by two comparisons, a Chi-squared test was used; if any groups contained less than five values, the Fisher’s Exact test was used in preference.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Galon J. Type, Density, and Location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006; 313:1960–4; PMID:17008531; <http://dx.doi.org/10.1126/science.1129139>
2. Galon J, Pagès F, Marincola FM, Angell HK, Thurin M, Lugli A, Zlobec I, Berger A, Bifulco C, Botti G et al. Cancer classification using the Immunoscore: a worldwide task force. *J Trans Med* 2012; 10:205; PMID:23034130; <http://dx.doi.org/10.1186/1479-5876-10-205>
3. Galon J, Pagès F, Marincola FM, Thurin M, Trinchieri G, Fox BA, Gajewski TF, Ascierto PA. The immune score as a new possible approach for the classification of cancer. *J Trans Med* 2012; 10:1; PMID:22214470; <http://dx.doi.org/10.1186/1479-5876-10-1>
4. Mlecnik B, Tosolini M, Charoentong P, Kirilovsky A, Bindea G, Berger A, Camus M, Gillard M, Bruneval P, Fridman WH et al. Biomolecular network reconstruction identifies T-Cell homing factors associated with survival in colorectal cancer. *Gastroenterology* 2010; 138:1429–40; PMID:19909745; <http://dx.doi.org/10.1053/j.gastro.2009.10.057>
5. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013; 39:782–95; PMID:24138885; <http://dx.doi.org/10.1016/j.immuni.2013.10.003>
6. Kim H, Nam SW, Rhee H, Shan Li L, Ju Kang H, Hye Koh K, Kyu Kim N, Song J, Tak-Bun Liu E, Kim H. Different gene expression profiles between microsatellite instability-high and microsatellite stable colorectal carcinomas. *Oncogene* 2004; 23:6218–25; PMID:15208663; <http://dx.doi.org/10.1038/sj.onc.1207853>
7. Phillips SM, Banerjee A, Feakins R, Li SR, Bustin SA, Dorudi S. Tumour-infiltrating lymphocytes in colorectal cancer with microsatellite instability are activated and cytotoxic. *British J Surg* 2004; 91:469–75; PMID:15048750; <http://dx.doi.org/10.1002/bjs.4472>
8. Buckowitz A, Knaebel HP, Benner A, Bläker H, Gebert J, Kienle P, von Knebel Doeberitz M, Kloor M. Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *British J Cancer* 2005; 92:1746–53; PMID:15856045; <http://dx.doi.org/10.1038/sj.bjc.6602534>
9. Banerjee A, Bustin SA, Dorudi S. The immunogenicity of colorectal cancers with high-degree microsatellite instability. *World J Surg Oncol* 2005; 3:26;

Acknowledgments

The authors thank the Colorectal Cancer Subtyping Consortium for kindly providing CMS classifications for the TCGA dataset.

Funding

NL is supported by a Cancer Research UK clinical PhD studentship. ADB is supported by a Wellcome Trust Postdoctoral Training fellowship for Clinician Scientists (ref 102732/Z/13/Z). BW is supported by a Wellcome Trust New Investigator Award (099266/Z/12/Z). GM is supported by Cancer Research UK.

Supplemental Material

Supplemental data for this article can be accessed on the publisher’s website.

- PMID:15890075; <http://dx.doi.org/10.1186/1477-7819-3-26>
10. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007; 50:113–30; PMID:17204026; <http://dx.doi.org/10.1111/j.1365-2559.2006.02549.x>
11. Fehlkner M, Huska MR, Jöns T, Andrade-Navarro MA, Kemmner W. Concerted down-regulation of immune-system related genes predicts metastasis in colorectal carcinoma. *BMC Cancer* 2014; 14:64; PMID:24495478; <http://dx.doi.org/10.1186/1471-2407-14-64>
12. Boissiere-Michot F, Lazennec G, Frugier H, Jarlier M, Roca L, Duffour J, Du Paty E, Laune D, Blanchard F, Le Pessot F et al. Characterization of an adaptive immune response in microsatellite-Instable colorectal cancer. *Oncolmmunology* 2014; 3:e29256; PMID:25101223; <http://dx.doi.org/10.4161/onci.29256>
13. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, Powderly JD, Carvajal RD, Sosman JA, Atkins MB et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012; 366:2443–54; PMID:22658127; <http://dx.doi.org/10.1056/NEJMoa1200690>
14. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, Drake CG, Camacho LH, Kauh J, Odunsi K et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 2012; 366:2455–65; PMID:22658128; <http://dx.doi.org/10.1056/NEJMoa1200694>
15. Willcox CR, Pitard V, Netzer S, Couzi L, Salim M, Silberzahn T, Moreau J-F, Hayday AC, Willcox BE, Déchanet-Merville J. Cytomegalovirus and tumor stress surveillance by binding of a human $\gamma\delta$ T cell antigen receptor to endothelial protein C receptor. *Nat Immunol* 2012; 13:872–9; PMID:22885985; <http://dx.doi.org/10.1038/ni.2394>
16. Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Salguero I et al. Germline mutations affecting the proof-reading domains of POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 2013; 45:136–44; PMID:23263490; <http://dx.doi.org/10.1038/ng.2503>
17. Dierness JW, de Miranda NF, Ferrone S, van Puijenbroek M, Cornelisse CJ, Fleuren GJ, van Wezel T, Morreau H. HNPCC versus sporadic microsatellite-unstable colon cancers follow different routes toward loss of HLA class I expression. *BMC Cancer* 2007; 7:33; PMID:17316446; <http://dx.doi.org/10.1186/1471-2407-7-33>
18. Dienstmann R, Guinney J, Delorenzi M, De Reynies A, Roepman P, Sadanandam A, Vermeulen L, Schlicker

- A, Missiaglia E, Soneson C et al. Colorectal Cancer Subtyping Consortium (CRCSC) identification of a consensus of molecular subtypes. *ASCO Meeting Abstracts* 2014; 32:3511.
19. Bennett SR, Carbone FR, Karamalis F, Miller JF, Heath WR. Induction of a CD8+ cytotoxic T lymphocyte response by cross-priming requires cognate CD4+ T cell help. *J Exp Med* 1997; 186:65–70; PMID:9206998; <http://dx.doi.org/10.1084/jem.186.1.65>
20. Bennett SR, Carbone FR, Karamalis F, Flavell RA, Miller JF, Heath WR. Help for cytotoxic-T-cell responses is mediated by CD40 signalling. *Nature* 1998; 393:478–80; PMID:9624004; <http://dx.doi.org/10.1038/30996>
21. Janssen EM, Lemmens EE, Wolfe T, Christen U, von Herrath MG, Schoenberger SP. CD4+ T cells are required for secondary expansion and memory in CD8+ T lymphocytes. *Nature* 2003; 421:852–6; PMID:12594515; <http://dx.doi.org/10.1038/nature01441>
22. Williams MA, Tyznik AJ, Bevan MJ. Interleukin-2 signals during priming are required for secondary expansion of CD8+ memory T cells. *Nature* 2006; 441:890–3; PMID:16778891; <http://dx.doi.org/10.1038/nature04790>
23. Xie Y, Akpınarli A, Maris C, Hipkiss EL, Lane M, Kwon EK, Muranski P, Restifo NP, Antony PA. Naive tumor-specific CD4(+) T cells differentiated in vivo eradicate established melanoma. *J Exp Med* 2010; 207:651–67; PMID:20156973; <http://dx.doi.org/10.1084/jem.20091921>
24. Quezada SA, Simpson TR, Peggs KS, Merghoub T, Vider J, Fan X, Blasberg R, Yagita H, Muranski P, Antony PA et al. Tumor-reactive CD4(+) T cells develop cytotoxic activity and eradicate large established melanoma after transfer into lymphopenic hosts. *J Exp Med* 2010; 207:637–50; PMID:20156971; <http://dx.doi.org/10.1084/jem.20091918>
25. Haabeth OA, Tveita AA, Fauskanger M, Schjesvold F, Lorvik KB, Hofgaard PO, Omholt H, Munthe LA, Dembic Z, Corthay A et al. How Do CD4(+) T Cells Detect and Eliminate Tumor Cells That Either Lack or Express MHC Class II Molecules? *Front Immunol* 2014; 5:174; PMID:24782871; <http://dx.doi.org/10.3389/fimmu.2014.00174>
26. Hunder NN, Wallen H, Cao J, Hendricks DW, Reilly JZ, Rodmyre R, Jungbluth A, Gnjatic S, Thompson JA, Yee C. Treatment of metastatic melanoma with autologous CD4+ T cells against NY-ESO-1. *N Engl J Med* 2008; 358:2698–703; PMID:18565862; <http://dx.doi.org/10.1056/NEJMoa0800251>

27. Tran E, Turcotte S, Gros A, Robbins PF, Lu YC, Dudley ME, Wunderlich JR, Somerville RP, Hogan K, Hinrichs CS et al. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science* 2014; 344:641-5; PMID:24812403; <http://dx.doi.org/10.1126/science.1251102>
28. Dong H, Strome SE, Salomao DR, Tamura H, Hirano F, Flies DB, Roche PC, Lu J, Zhu G, Tamada K et al. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat Med* 2002; 8:793-800; PMID:12091876; <http://dx.doi.org/10.1038/nm0902-1039c>
29. Lee SJ, Jang BC, Lee SW, Yang YI, Suh SI, Park YM, Oh S, Shin JG, Yao S, Chen L et al. Interferon regulatory factor-1 is prerequisite to the constitutive expression and IFN-gamma-induced upregulation of B7-H1 (CD274). *FEBS Lett* 2006; 580:755-62; PMID:16413538; <http://dx.doi.org/10.1016/j.febslet.2005.12.093>
30. Woo SR, Turnis ME, Goldberg MV, Bankoti J, Selby M, Nirschl CJ, Bettini ML, Gravano DM, Vogel P, Liu CL et al. Immune inhibitory molecules LAG-3 and PD-1 synergistically regulate T-cell function to promote tumoral immune escape. *Cancer Res* 2011; 72:917-27; PMID:22186141; <http://dx.doi.org/10.1158/0008-5472.CAN-11-1620>
31. Løvig T, Andersen SN, Thorstensen L, Diep CB, Meling GI, Lothe RA, Rognum TO. Strong HLA-DR expression in microsatellite stable carcinomas of the large bowel is associated with good prognosis. *British J Cancer* 2002; 87:756-62; PMID:12232760; <http://dx.doi.org/10.1038/sj.bjc.6600507>
32. Saeterdal I, BJORHEIM J, LISLERUD K, GJERTSEN MK, BUKHOLM IK, OLSEN OC, NESLAND JM, ERIKSEN JA, MOLLER M, LINDBLOM A et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proc Natl Acad Sci U S A* 2001; 98:13255-60; PMID:11687624; <http://dx.doi.org/10.1073/pnas.231326898>
33. Noshu K, Baba Y, Tanaka N, Shima K, Hayashi M, Meyerhardt JA, Giovannucci E, Dranoff G, Fuchs CS, Ogino S. Tumor-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. *J Pathol* 2010; 222:350-66; PMID:20927778; <http://dx.doi.org/10.1002/path.2774>
34. Maudsley DJ, Morris AG. Kirsten murine sarcoma virus abolishes interferon gamma-induced class II but not class I major histocompatibility antigen expression in a murine fibroblast line. *J Exp Med* 1988; 167:706-11; PMID:2831293; <http://dx.doi.org/10.1084/jem.167.2.706>
35. Maudsley DJ, Bateman WJ, Morris AG. Reduced stimulation of helper T cells by Ki-ras transformed cells. *Immunology* 1991; 72:277-81; PMID:1826672
36. Atkins D, Breuckmann A, Schmahl GE, Binner P, Ferrone S, Krummenauer F, Störkel S, Seliger B. MHC class I antigen processing pathway defects, ras mutations and disease stage in colorectal carcinoma. *Int J Cancer* 2004; 109:265-73; <http://dx.doi.org/10.1002/ijc.11681>
37. Kortylewski M, Yu H. Role of Stat3 in suppressing anti-tumor immunity. *Curr Opin Immunol* 2008; 20:228-33; PMID:18479894; <http://dx.doi.org/10.1016/j.coi.2008.03.010>
38. Melief CJ. Cancer immunotherapy by dendritic cells. *Immunity* 2008; 29:372-83; PMID:18799145; <http://dx.doi.org/10.1016/j.immuni.2008.08.004>
39. Hutti JE, Pfeifferle AD, Russell SC, Sircar M, Perou CM, Baldwin AS. Oncogenic PI3K mutations lead to NF-kappaB-dependent cytokine expression following growth factor deprivation. *Cancer Res* 2012; 72:3260-9; PMID:22552288; <http://dx.doi.org/10.1158/0008-5472.CAN-11-4141>
40. Thun MJ, Namboodiri MM, Heath CW, Jr. Aspirin use and reduced risk of fatal colon cancer. *N Engl J Med* 1991; 325:1593-6; PMID:1669840; <http://dx.doi.org/10.1056/NEJM199112053252301>
41. Liao X, Lochhead P, Nishihara R, Morikawa T, Kuchiba A, Yamauchi M, Imamura Y, Qian ZR, Baba Y, Shima K et al. Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. *New Eng J Med* 2012; 367:1596-606; PMID:23094721; <http://dx.doi.org/10.1056/NEJMoa1207756>
42. Domingo E, Church DN, Sieber O, Ramamoorthy R, Yanagisawa Y, Johnstone E, Davidson B, Kerr DJ, Tomlinson IP, Midgley R. Evaluation of PIK3CA mutation as a predictor of benefit from nonsteroidal anti-inflammatory drug therapy in colorectal cancer. *J Clin Oncol* 2013; 31:4297-305; PMID:24062397; <http://dx.doi.org/10.1200/JCO.2013.50.0322>
43. Cbioportal for Cancer Genomics. Memorial Sloan Kettering Cancer Center.
44. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; 6:p11.
45. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; 2:401-4; PMID:22588877; <http://dx.doi.org/10.1158/2159-8290.CD-12-0095>
46. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; 487:330-7; PMID:22810696; <http://dx.doi.org/10.1038/nature11252>