

Tumour gene expression predicts response to cetuximab in patients with *KRAS* wild-type metastatic colorectal cancer

JB Baker^{*,1}, D Dutta¹, D Watson¹, T Maddala¹, BM Munneke¹, S Shak¹, EK Rowinsky², L-A Xu³, CT Harbison³, EA Clark³, DJ Mauro³ and S Khambata-Ford³

¹Genomic Health, Inc., 301 Penobscot Drive, Redwood City, CA, USA; ²ImClone Systems, New York, NY, USA; ³Bristol-Myers Squibb, Princeton, NJ, USA

BACKGROUND: Although it is accepted that metastatic colorectal cancers (mCRCs) that carry activating mutations in *KRAS* are unresponsive to anti-epidermal growth factor receptor (EGFR) monoclonal antibodies, a significant fraction of *KRAS* wild-type (wt) mCRCs are also unresponsive to anti-EGFR therapy. Genes encoding EGFR ligands amphiregulin (*AREG*) and epiregulin (*EREG*) are promising gene expression-based markers but have not been incorporated into a test to dichotomise *KRAS* wt mCRC patients with respect to sensitivity to anti-EGFR treatment.

METHODS: We used RT–PCR to test 110 candidate gene expression markers in primary tumours from 144 *KRAS* wt mCRC patients who received monotherapy with the anti-EGFR antibody cetuximab. Results were correlated with multiple clinical endpoints: disease control, objective response, and progression-free survival (PFS).

RESULTS: Expression of many of the tested candidate genes, including *EREG* and *AREG*, strongly associate with all clinical endpoints. Using multivariate analysis with two-layer five-fold cross-validation, we constructed a four-gene predictive classifier. Strikingly, patients below the classifier cutpoint had PFS and disease control rates similar to those of patients with *KRAS* mutant mCRC.

CONCLUSION: Gene expression appears to identify *KRAS* wt mCRC patients who receive little benefit from cetuximab. It will be important to test this model in an independent validation study.

British Journal of Cancer (2011) **104**, 488–495. doi:10.1038/sj.bjc.6606054 www.bjancer.com

Published online 4 January 2011

© 2011 Cancer Research UK

Keywords: EGFR; colorectal cancer; cetuximab

The epidermal growth factor receptor (EGFR) has been implicated in the growth and aggressiveness of a number of different cancers (Mendelsohn and Baselga, 2006), several of which are responsive to drugs that target this receptor (Baselga and Arteaga, 2005; Modjtahedi and Essapen, 2009). The anti-EGFR monoclonal antibody cetuximab has demonstrated clinical benefit in, and is widely used to treat metastatic colorectal cancer (mCRC) (Jonker *et al*, 2009; Van Cutsem *et al*, 2010). Efforts have been made to identify biomarkers that will optimise patient selection for treatment to maximise the therapeutic index for mCRC patients receiving cetuximab and other anti-EGFR therapies.

Activating *KRAS* mutations, which are present in 30–40% of CRC (Samowitz *et al*, 2000; Andreyev *et al*, 2001) serve as a key useful biomarker in this context. A number of independent studies strongly link the presence of a somatic activating mutation in the *KRAS* gene (codons 12 or 13) with nearly complete mCRC resistance to anti-EGFR monoclonal antibodies cetuximab (Benvenuti *et al*, 2007; Di Fiore *et al*, 2007; Khambata-Ford *et al*, 2007; De Roock *et al*, 2008; Karapetis *et al*, 2008; Lièvre *et al*, 2008) and panitumumab (Benvenuti *et al*, 2007; Amado *et al*, 2008; Freeman *et al*, 2008), leading the American Society of Clinical Oncology to recommend that only mCRC patients with *KRAS* wild-type (wt) mCRC be considered candidates to receive anti-EGFR therapy

(Allegra *et al*, 2009). Other investigated mutations in mCRC are either rare (e.g., activating EGFR mutations (Barber *et al*, 2004; Tsuchihashi *et al*, 2005)) or, as a result of conflicting data, controversial with respect to response to anti-EGFR monoclonal antibodies (e.g., *PI3K* and *BRAF*); (Samowitz *et al*, 2005; Di Nicolantonio *et al*, 2008; Loupakis *et al*, 2009; Perrone *et al*, 2009; Prenen *et al*, 2009; Sartore-Bianchi *et al*, 2009; Souglakos *et al*, 2009; Roth *et al*, 2010). Despite the significance of *KRAS* mutations, the efficacy of anti-EGFR monoclonal antibodies in the 60–70% of mCRC patients with *KRAS* wt tumours is still limited, with response rates between 10 and 40% (Allegra *et al*, 2009). There is a need for additional predictive biomarkers for these patients. Interestingly, the expression of the EGFR protein has not been strongly associated with clinical response to cetuximab in colorectal cancer (Cunningham *et al*, 2004; Saltz *et al*, 2004; Chung *et al*, 2005; Meropol, 2005), although there is limited evidence that amplification of the *EGFR* gene relates to objective response and other indices of clinical benefit (Moroni *et al*, 2005; Cappuzzo *et al*, 2008; Personeni *et al*, 2008).

Two independent groups recently reported that increased expression of genes encoding two EGFR ligands, amphiregulin (*AREG*) and epiregulin (*EREG*) strongly associates with increased therapeutic benefit from cetuximab in mCRC patients (Khambata-Ford *et al*, 2007; Jacobs *et al*, 2009; Jonker *et al*, 2009). Although these results are encouraging, a gene expression-based test with adequate performance has not yet been developed. An expanded model with markers additional to *AREG* and *EREG* may be

*Correspondence: Dr JB Baker; E-mail: j baker@genomichealth.com
Received 30 June 2010; revised 16 November 2010; accepted 19 November 2010; published online 4 January 2011

optimal. For practical clinical utility the test would categorise patients for treatment, that is, have dichotomising cutpoints (Jacobs *et al*, 2009).

With the goal of creating a clinically useful test, we have used high-throughput RT-PCR to explore a set of 110 biologically based candidate gene expression biomarkers in formalin-fixed, paraffin-embedded primary tumour specimens from 144 *KRAS* wt mCRC patients who had been treated with cetuximab monotherapy. There is a strong rationale for this strategy of testing candidate genes based on the known biology of the EGFR pathway as evidenced by the above referenced association of *KRAS* to drug resistance and *AREG* and *EREG* to drug sensitivity. The fact that constitutive activation of *KRAS* marks resistance to cetuximab is consistent with the known role of *KRAS* as a key downstream signal transducer of the EGFR pathway (Mendelsohn and Baselga, 2006). The findings that increased *AREG* and *EREG* expression associate with sensitivity to cetuximab accords with the general concept of oncogene addiction (Weinstein, 2002).

MATERIALS AND METHODS

Colorectal cancer tissue samples were obtained from the primary colon tumours removed at initial surgical resection for patients from three cetuximab monotherapy studies: IMC CP02-0144, IMC CP02-0141, and BMS CA225-045. None of the tissue samples were obtained from metastases and a single sample was obtained from the primary tumours. Eligibility in the IMC CP02-0141 study required that patients had previous therapy with at least one chemotherapeutic regimen for advanced disease that included a fluoropyrimidine and irinotecan, and documentation that their disease was refractory to this treatment. Patients were eligible for IMC CP02-0144 and BMS CA225-045 if their disease was refractory to irinotecan-, oxaliplatin-, and fluoropyrimidine-based regimens. In all three studies, eligibility criteria also included an Eastern Cooperative Oncology Group performance status of two or less. In addition, patients were required to be at least 18 years of age; patients in IMC CP02-0144 and BMS CA225-045 could not have received major surgery, radiation chemotherapy, or investigational agents within 4 weeks. Study CP02-0141 also indicated that no surgery was permitted within 21 days of study entry, excluding previous diagnostic biopsy. Standard laboratory test values had to be within normal limits. Patients were treated with cetuximab at a loading dose of 400 mg m⁻² followed by 250 mg m⁻² weekly. In one study, patients could escalate their dose after the first 3 weeks and every 3 weeks thereafter to a maximum dose of 400 mg m⁻², if they did not experience a skin rash greater than grade 2. All protocols were approved by institutional review boards at the participating institutions.

RNA was extracted from a minimum of three to six 5- μ m unstained sections of FPE tumour tissue on glass slides without a cover slip, using protocols optimised for colon FPE tissue by Genomic Health Inc. (Redwood City, CA, USA). One slide from each patient was H&E stained and used as a guide for removal of non-tumour tissue by manual microdissection. RNA was isolated and TaqMan RT-PCR was carried out as previously described (Clark-Langone *et al*, 2007).

Candidate genes (shown in Supplementary Table S1) were selected for gene expression analysis, based on previous evidence that they may be linked to the sensitivity or resistance to EGFR-inhibitor drugs either directly (as components of the EGFR signaling pathway) or indirectly (by marking fundamental pathobiological variables such as cell cycle activity, or stromal response). As shown in Supplementary Figure S1, using the initial set of 254 samples, expression levels of 99 genes were evaluated. Both emerging evidence regarding the importance of *KRAS* mutation status in response to EGFR inhibitors and the desire to investigate additional cancer-related genes as well as alternative

reference genes resulted in an optimised version of the assay gene panel. Genes significantly associated with outcome in the original assay were carried forward to the optimised assay. Of the original 254 samples, 226 (those with sufficient RNA remaining) were re-assayed with the updated panel of 57 genes plus *KRAS* sequences. All together, 110 genes were evaluated.

A normalisation method was used to compensate for systematic sources of variation in gene expression measurements. In the optimised version of the assay, the cycle threshold (C_T) measurement for each test gene was normalised relative to the average expression of five reference genes (*UBB*, *ATP5E*, *PGK1*, *VDAC2*, and *GPX1*) in that tumour RNA sample. The reference genes were selected for this purpose in the course of RT-PCR studies of several hundred genes carried out in over 1800 stage II and III colorectal cancer specimens. For details regarding the assay formats, see Supplementary Figure S1.

Wild-type and mutant *KRAS* were detected using allele-specific blocker PCR that uses a PCR primer that is specific for the mutant template, and an oligonucleotide that blocks amplification of the wt gene (Morlan *et al*, 2009). Samples were assayed for the seven most frequent mutations reported in *KRAS* codons 12 and 13 (Gly12Asp, Gly12Ala, Gly12Val, Gly12Ser, Gly12Arg, Gly12Cys, and Gly13Asp), accounting for approximately 97% of all *KRAS* mutations observed in CRC.

Clinical endpoints and statistical methods

Complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD) was assessed by study investigators. Complete response was defined as complete disappearance of all tumour lesions for at least 4 weeks from the date of documentation of CR. Partial response was defined as a decrease (relative to baseline) of 50% or greater in the sum of the products of the two largest perpendicular diameters of all index lesions, as determined by two observations at least 4 weeks apart. Additionally, patients could not meet the criteria for PD. Progressive disease was defined as at least a 25% increase in the sum of the products of all index lesions (taking as reference the smallest sum recorded at or following baseline) or progression of non-measurable lesions or the appearance of any new lesions. Stable disease was defined as failure to meet criteria for CR or PR in the absence of PD. Disease control was defined as having either CR, PR, or SD as a patient's best response. Objective response rate (ORR) was defined as the proportion of patients achieving a CR or PR as a patient's best response. Progression-free survival (PFS) was defined as the time from entry onto the study to the first date of progression for patients with documented PD, or the date of death for patients who died within 90 days of their last tumour assessment without documented progression. Patients who did not progress or die were censored at the date of their last tumour assessment. Gene expression was standardised by dividing each gene's reference-normalised expression level (on the log₂ C_T scale) by its sample s.d. Univariate logistic regression models were used to examine the associations between the expression of test genes and outcome (disease control and objective response). Univariate Cox proportional hazards regression models were used to examine the associations between the expression of test genes or the four-gene score and PFS. For each regression model, the *P*-value under the likelihood ratio test of the null hypothesis that the odds ratio (or hazard ratio) equals one is reported. All statistical tests are two-sided and *P* < 0.05 is considered significant. The minimum false discovery rate (FDR) at which each null hypothesis could be rejected using the method of Benjamini and Hochberg (1995) is also provided as a conservative statistical adjustment for the multiplicity of testing multiple gene models.

KRAS expression measurements were used to classify each patient as either positive or negative for at least one *KRAS*

mutation. Using a logistic regression model, a probability of mutation for each *KRAS* mutation sequence and sample was estimated. When the predicted probability of a mutation was >0.5 , the sample was classified as positive for the particular *KRAS* mutation variant. The logistic regression model applied to these data were developed using samples from an external training data set in which the *KRAS* mutation status of each sample was known. *P*-values comparing the proportion of responses (disease control or objective response) in patients with and without *KRAS* mutations were calculated using the χ^2 test.

Multi-gene classification models were developed using two-layer, five-fold cross-validation approaches in the *KRAS* wt subset of patients (Ambrose and McLachlan, 2002). A supervised approach was used, analysing only the subset of genes that were

Table 1 *KRAS* wild-type patients by endpoint and study

Study	Disease control endpoint		Objective response endpoint		Total
	Responder	Non-responder	Responder	Non-responder	
CA 225045	12	11	5	18	23
IMCL CP02-0141	8	3	3	8	11
IMCL CP02-0144	67	43	23	87	110
Total	87 ^a	57 ^b	31 ^c	113 ^d	144

Abbreviations: CR = complete response; PD = progressive disease; PR = partial response; SD = stable disease. ^aIncludes 1 CR, 30 PR, and 56 SD. ^bIncludes PD only. ^cIncludes 1 CR and 30 PR. ^dIncludes 57 PD and 56 SD.

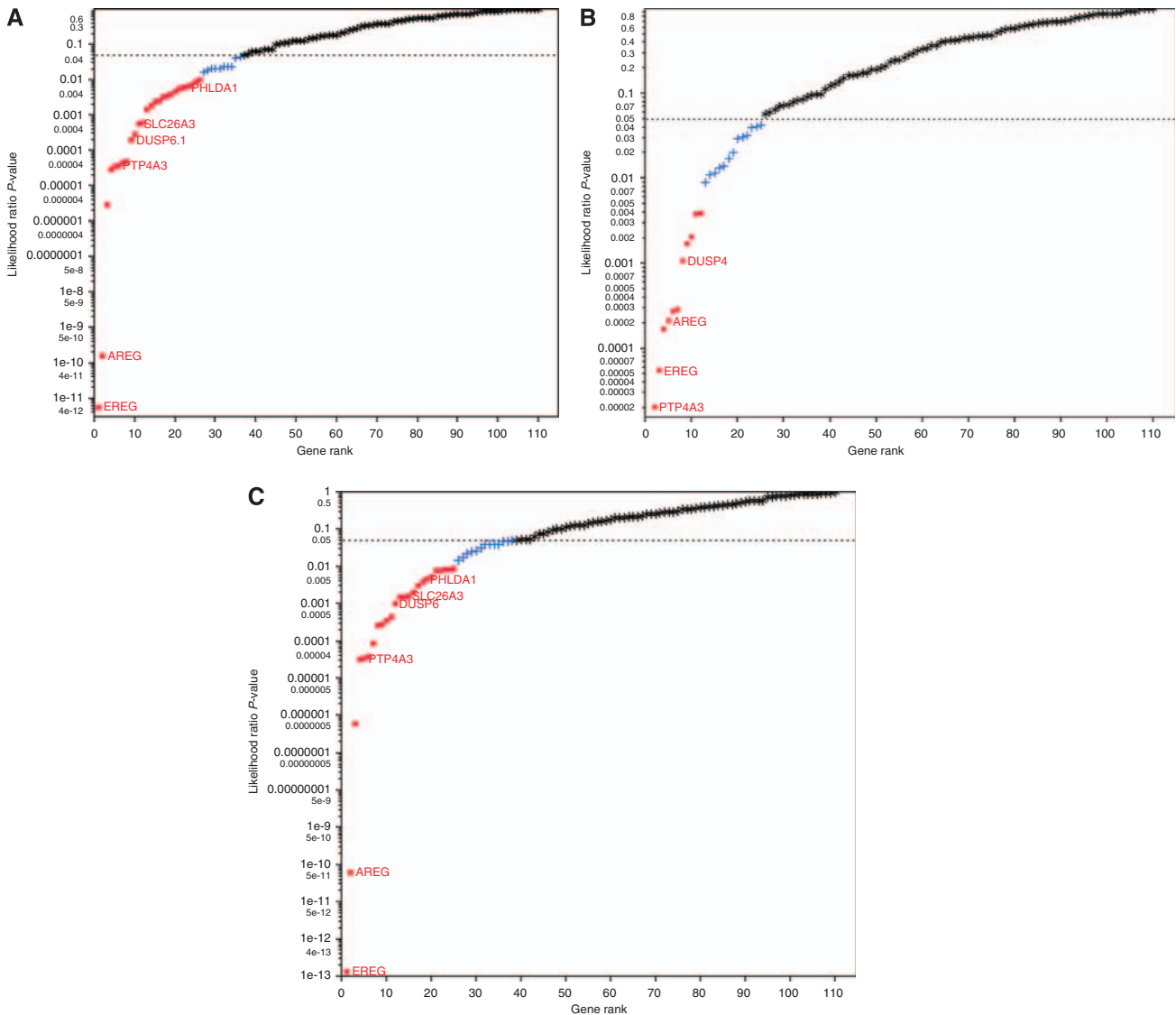


Figure 1 (A) Distribution of likelihood ratio *P*-values in *KRAS* wt patients by gene rank (disease control). Note: ■ and + = genes significantly associated with outcome, unadjusted *P*-value <0.05 , ■ = genes associated with outcome controlling for FDR <0.05 , * = genes not significantly associated with outcome. In total, 110 genes depicted. (B) Distribution of likelihood ratio *P*-values in *KRAS* wt patients by gene rank (objective response). Note: ■ and + = genes significantly associated with outcome, unadjusted *P*-value <0.05 , ■ = genes associated with outcome controlling for FDR <0.05 , * = genes not significantly associated with outcome. In total, 110 genes depicted. (C) Distribution of likelihood ratio *P*-values in *KRAS* wt patients by gene rank (PFS). Note: ■ and + = genes significantly associated with outcome, unadjusted *P*-value <0.05 , ■ = genes associated with outcome controlling for FDR <0.05 , * = genes not significantly associated with outcome. In total, 110 genes depicted.

identified as significantly associated with outcome based on univariate analyses of gene expression and outcome (disease control and objective response). Forward stepwise logistic regression was employed in the inner layers and inspection of the area under the curve (AUC) for each of the models was used to identify the optimal multi-gene model. The outer layers were used to evaluate the models constructed in the inner layers. The two-layer, five-fold cross-validation was repeated 50 times using different partitions of the data, yielding 250 models. The threshold for the final four-gene model was determined by identifying the cutpoint that yielded the highest number of true negatives plus true positives on the disease control endpoint. Median time to progression and/or death was calculated using Kaplan–Meier (product-limit estimation) methods and associated 95% confidence intervals were calculated based on the sign test (Brookmeyer and Crowley, 1982). Corresponding Kaplan–Meier plots were used to display PFS for the selected four-gene classifier on the disease control endpoint. The log-rank test was calculated to test the null hypothesis that there is no difference in the probability of progression and/or death over the entire time period among the groups identified by the multi-gene classifier.

RESULTS

Patient characteristics

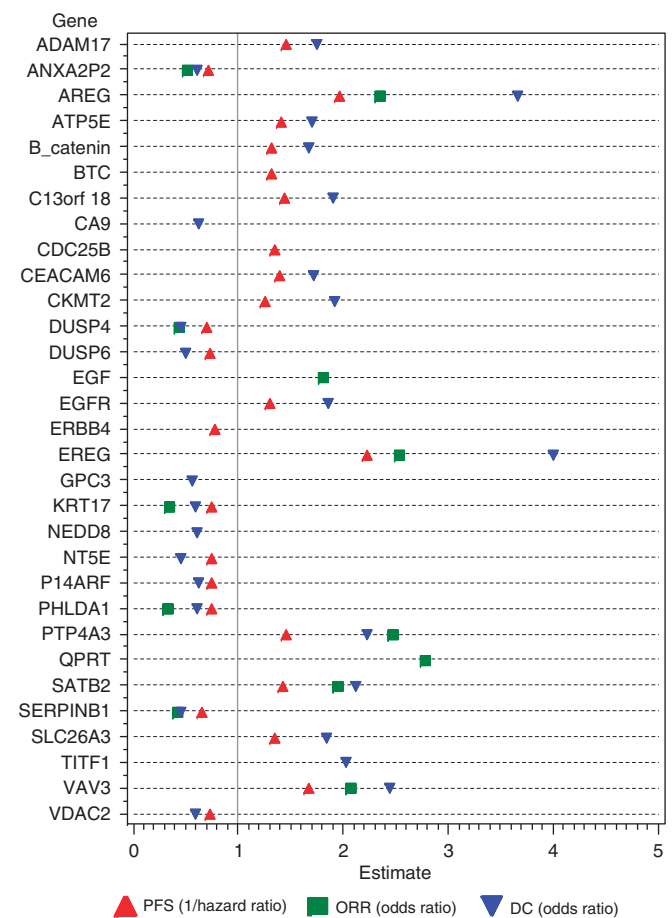
A total of 326 tumour specimens were obtained from the primary colon tumours removed at initial surgical resection for patients who were enrolled in cetuximab monotherapy trials. After exclusion of 46 specimens based on predefined pathology and laboratory criteria (10 because of insufficient tumour, 22 because of insufficient RNA, and 14 because of laboratory failure), 280 patients had successful gene expression analysis. An additional 26 patients were clinically not evaluable (not treated, or withdrew or died before tumour assessment), resulting in 254 patients that were evaluable for the RT-PCR-based gene expression analysis shown in Supplementary Figure S1. There were 226 samples with sufficient RNA for both (*KRAS*) mutation and gene expression analysis (174 patients from IMC CP02-0144, 13 patients from IMC CP02-0141, and 39 patients from BMS CA225-045).

An activating *KRAS* (codons 12 or 13) mutation was identified in 82 (36%) of the 226 patients. Both this frequency, as well as the distribution of the individual codon 12 and 13 mutations (Supplementary Table S2), are in accord with the published literature (Samowitz *et al*, 2000). Also, consistent with published literature, the disease control, ORR and PFS rates were very low in patients with *KRAS* mutant compared with *KRAS* wt tumours (Supplementary Table S3 and Supplementary Figure S2). The 144 patients with *KRAS* wt tumours were carried forward for gene expression analysis.

Of the 144 patients, 1 patient (<1%) had a CR, 30 patients (21%) had a PR, 56 patients (39%) had SD, and 57 (40%) had PD (Table 1).

Gene expression and sensitivity to cetuximab in *KRAS* wt mCRC

Among the 144 patients with wt *KRAS* tumours, the expression levels of 26, 12, and 25 genes were associated with the disease control, objective response, and PFS endpoints, respectively (FDR <0.05 after adjusting for multiplicity). As shown by Figures 1A–C, many of these gene expression values had highly significant *P*-values. Further, as shown in Figure 2, many of these genes are associated with outcome over multiple endpoints (nine of them across all three endpoints). These genes include *DUSP6*, which encodes a dual specificity phosphatase, and *EREG* and *AREG*, which encode the EGFR ligands epiregulin and



Note: hazard ratios and odds ratios represent the effect for one standard deviation change in gene expression.

Figure 2 Genes associated with outcome after adjusting for multiplicity across all three endpoints in *KRAS* wt patients.

amphiregulin, respectively. Increased expression of *DUSP6*, which blocks the EGFR pathway at MAPK, is associated with decreased likelihood of disease control, while increased expression of *AREG* and *EREG*, which stimulate the EGFR pathway, are associated with increased likelihood of disease control. Although *EREG* and *AREG* tightly co-expressed, they did not co-express with other significant genes, consistent with the possibility that the other genes carry independent predictive information (see gene expression dendrogram in Supplementary Figure S3).

Unlike *AREG* and *EREG*, genes encoding other EGFR ligands, TGFalpha, HB-EGF, and EGF, were not consistently associated with outcome. Our results indicate that *AREG* and *EREG* tend to be more highly expressed than these other ligands in mCRC (see Supplementary Figure S4).

It should be noted that not all of the highly significant genes are known to be linked to the EGFR pathway. For example, *PTP4A3*, a gene encoding a protein tyrosine phosphatase that has been reported to be expressed at increased levels in aggressive (Mollevi *et al*, 2008) or metastatic (Guzinska-Ustymowicz *et al*, 2009) CRC, strongly associated with therapy benefit across all three endpoints (Figure 2).

Of the three studies combined for this analysis, study IMCL CP02-0144 accounted for 76% of the total patients (Table 1). Results in the subset of patients from study IMCL CP02-0144 alone were highly consistent with results obtained when analysing all three studies combined (data not shown).

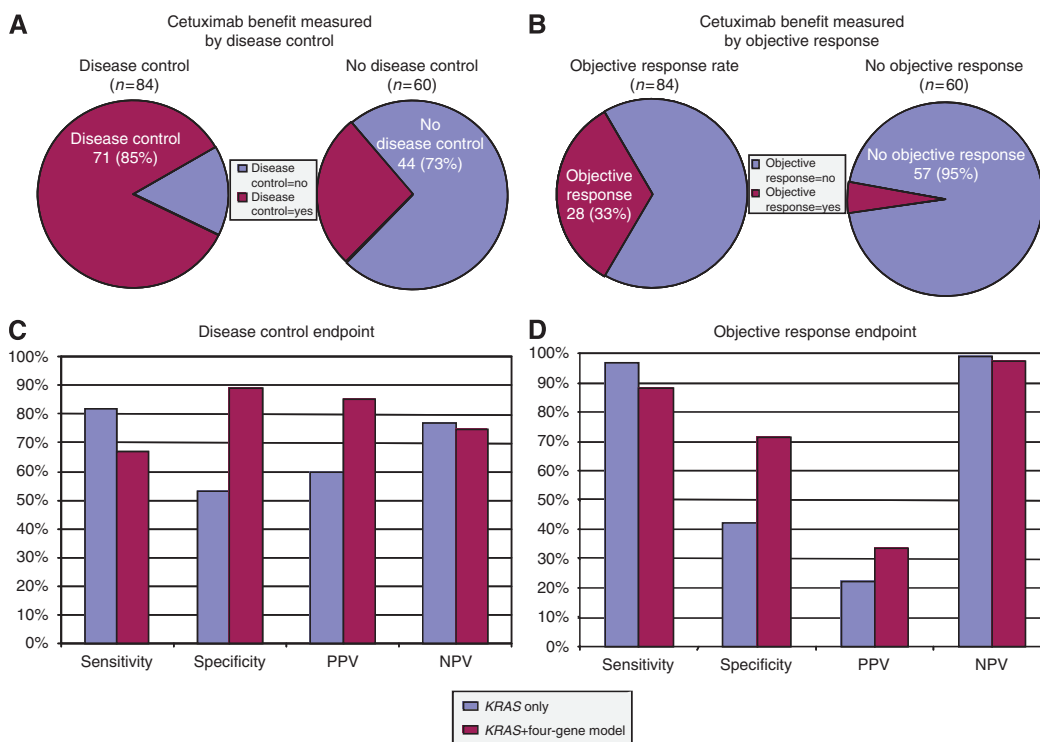


Figure 3 (A) Four-gene classification model for prediction of cetuximab benefit measured by disease control rate in *KRAS* wt patients. (B) Four-gene classification model for prediction of cetuximab benefit measured by objective response rate in *KRAS* wt patients. (C) Performance measures of two patient classifiers (disease control endpoint). (D) Performance measures of two patient classifiers (objective response endpoint). The four-gene classifier includes *DUSP6*, *SLC23A2* and the average of *EREG* and *AREG*. *EREG* and *AREG* were averaged because they tightly co-express in these colon cancer specimens ($R^2 = 0.83$).

Multi-gene classifier developed in patients with *KRAS* wt mCRC

In view of the substantial fraction of tested genes whose expression was significantly associated with each of the clinical endpoints, we explored the potential performance of multi-gene classifiers. Applying two-layer five-fold cross-validation in patients with *KRAS* wt tumours, we identified several genes that consistently performed well as a group based on diagnostic measures (AUC, positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity) in the outer layer of the cross-validation. We considered models with differing numbers of genes, including a two-gene model with *AREG* and *EREG* only. Evaluation of receiver operating characteristic curves indicated that a two-gene model has a sharper shoulder than a four-gene model, which implies that this model is more sensitive to the choice of a cutpoint than the four-gene model. The additional robustness of the four-gene model to choice of cutpoint, combined with the desire to minimise analytical variation suggested that a four-gene model was optimal. A final model that yields a predictive score includes four genes: *EREG*, *AREG*, and two other genes, *DUSP6* (which encodes a dual-specificity phosphatase) and *SLC26A3* (which encodes an intestinal chloride ion transporter that is also known as downregulated in colon adenoma).

A cutpoint in the four-gene score was selected to construct a classifier that identifies *KRAS* wt patients most likely to benefit from cetuximab treatment. The cutpoint for the four-gene model is the maximum of the sum of the true negative and true positive values for the disease control endpoint when the predictive score was applied. As shown in Figure 3A, this selected cutpoint would lead to treatment of 84 patients (58% of *KRAS* wt patients) out of which 71 would have disease control (85% of these treated patients). Sixty patients would not have been treated, 27% of whom would have disease control (the majority (82%) of whom had SD).

When the four-gene score with the same cutpoint is applied to objective response, 84 patients would be treated, and 28 (33%) would respond; 60 patients would not be treated, and 3 of these (5%) would respond (Figure 3B). With respect to both the disease control and objective response endpoints, use of this multi-gene classifier in addition to *KRAS* status significantly improves specificity and PPV when compared with a patient selection based on *KRAS* status alone, but with a slight decrease in sensitivity (Figure 3C and D).

The performance of the four-gene classifier with respect to PFS is shown in Figure 4. The median PFS for *KRAS* wt patients with a four-gene score at or above the cutpoint is 163 days compared with a median PFS for all *KRAS* wt patients of 103 days. It is noteworthy that the 60 *KRAS* wt patients with a four-gene score below the cutpoint had a median time to progression and/or death similar to the 82 patients whose tumours carried a *KRAS* mutation (40 and 41 days, respectively; hazard ratio for this comparison is 1.04 (95% CI: 0.74–1.46)). The hazard ratio for comparison of *KRAS* wt patients above and below the four-gene score cutpoint is 0.22 (95% CI: 0.15–0.32). A test of the proportional hazards assumption indicated that the assumption was not violated ($P = 0.41$). However, it should be stated that the performance of the four-gene classifier on the PFS endpoint is evaluated here in the same group of patients that was used to develop the classifier based on objective response and disease control endpoints.

DISCUSSION

Here, we report that a number of candidate gene expression markers, many of which were selected based on their relationship to the EGFR pathway, strongly associate with disease control, PFS and ORR in *KRAS* wt mCRC patients receiving monotherapy with the anti-EGFR antibody cetuximab. Multivariate analysis

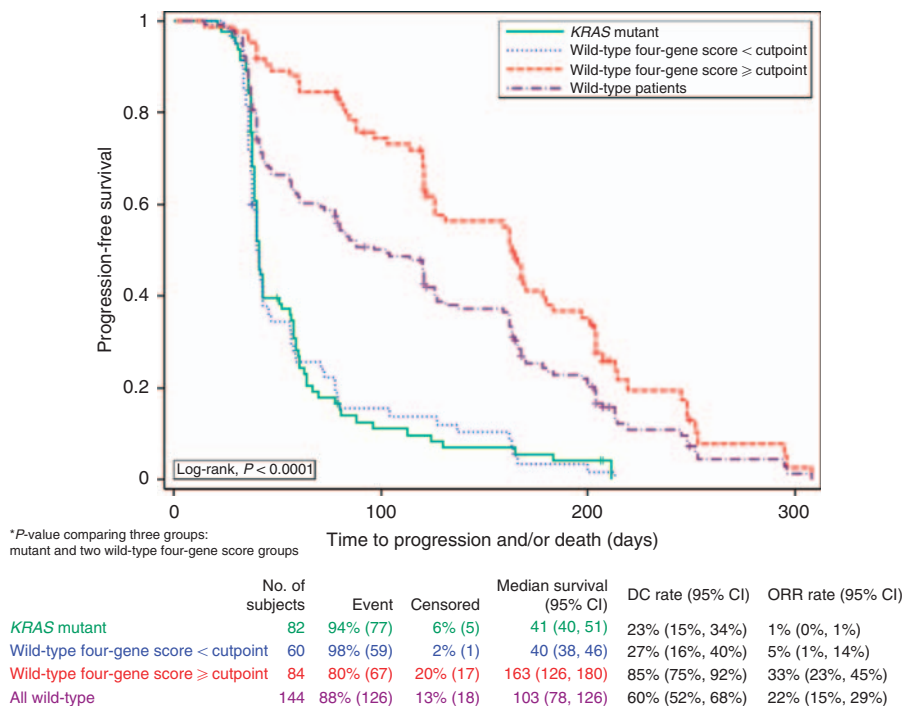


Figure 4 Progression-free survival in patients classified using *KRAS* and a four-gene classifier.

evaluating gene expression associated with disease control yields a four-gene model that predicts the likelihood that individual patients with *KRAS* wt mCRC will respond to cetuximab. This model sharply dichotomises patients with respect to likelihood of response, which has practical potential in the clinical setting. We propose that this model is sufficiently promising that it should be brought forward for testing in a clinical validation study.

Two of the genes in this model, *AREG* and *EREG*, have been previously identified as strong markers of sensitivity to anti-EGFR antibodies: by one study that analysed metastatic mCRC tissue from patients receiving cetuximab monotherapy (Khambata-Ford et al, 2007), and by another study that analysed primary mCRC tissue from patients receiving cetuximab in combination with chemotherapy (Jacobs et al, 2009; Jonker et al, 2009). Although *AREG* and *EREG* have not been studied in a large validation trial with a non-treated patient arm, our results with those of the two previous studies provide substantial evidence that these two markers are real and robust.

To develop an optimised gene expression-based test, we have examined additional candidate gene expression markers. The discoveries that key regulatory genes in the EGFR pathway, in the form of *KRAS*, *EREG*, and *AREG*, are anti-EGFR predictive markers suggested to us the potential value of screening other candidate markers based on both the biology of this pathway as well as other features of the pathobiology of CRC. As this study examines primary mCRC tissue from patients receiving cetuximab monotherapy, it is unencumbered by the potentially confounding complication presented by combination therapy. Further, it incorporates both a relatively large number of *KRAS* wt patients and multiple clinical endpoints.

The strength of association of many of these markers with clinical endpoints is remarkable: nine genes associating with all three endpoints (FDR < 0.05) after adjusting for multiple comparisons. Although further work needs to be done to confirm the performance of the identified markers, a number of them are compatible with known biology, in particular, the concept of oncogene addiction (Weinstein, 2002). Thus, *AREG* and *EREG*,

which encode activators of the EGFR pathway, associate with increased likelihood of clinical benefit; *DUSP6*, which encodes a phosphatase inhibitor of the EGFR pathway, associates with decreased likelihood of clinical benefit.

Although this study does not include a control arm of untreated patients, and therefore leaves open the possibility that a number of the described biomarkers are prognostic rather than predictive of cetuximab benefit, it seems probable that many of the top genes identified are in fact predictive of drug response because the objective response endpoint lends itself to discovery of predictive rather than prognostic biomarkers. An objective response in colon cancer is generally recognised as a treatment effect.

We performed multivariate analysis on the entire set of identified markers to generate a four-gene classifier that, in exploratory analysis, performs better than *EREG/AREG*. Used in conjunction with *KRAS* mutation status, this assay associates with several indices of clinical outcome better than either gene expression alone or *KRAS* mutation status alone.

The four-gene classifier is exploratory, as this model was discovered and estimated using the same *KRAS* wt patients. The degree of over-fitting was assessed in part through two-layer, five-fold cross-validation. External validation of this multi-gene model will require an independent study that is adequately statistically powered and includes a control arm of patients who are not treated with cetuximab.

Our findings are based on analysis of formalin-fixed primary tumour tissue, which is generally obtained at the time of definitive colorectal tumour resection. Therefore, if this classifier were to be validated, it will be technically feasible to introduce it into clinical practice. The applicability of these results to patients treated with cetuximab in combination therapy deserves further study.

ACKNOWLEDGEMENTS

Part of this study was presented at the 2008 ASCO Annual Meeting (JB Baker, presenter).

Conflict of interest

The authors are employees of the respective companies listed.

REFERENCES

- Allegra CJ, Jessup JM, Somerfield MR, Hamilton SR, Hammond EH, Hayes DF, McAllister PK, Morton RF, Schilsky RL (2009) American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy. *J Clin Oncol* **27**: 2091–2096
- Amado RG, Wolf M, Peeters M, Van Cutsem E, Siena S, Freeman DJ, Juan T, Sikorski R, Suggs S, Radinsky R, Patterson SD, Chang DD (2008) Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* **26**: 1626–1634
- Ambroise C, McLachlan GJ (2002) Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci USA* **99**: 6562–6566
- Andreyev HJ, Norman AR, Cunningham D, Oates J, Dix BR, Iacopetta BJ, Young J, Walsh T, Ward R, Hawkins N, Beranek M, Jandik P, Benamouzig R, Jullian E, Laurent-Puig P, Olschwang S, Muller O, Hoffmann I, Rabes HM, Zietz C, Troungos C, Valavanis C, Yuen ST, Ho JW, Croke CT, O'Donoghue DP, Giarretti W, Rapallo A, Russo A, Bazan V, Tanaka M, Omura K, Azuma T, Ohkusa T, Fujimori T, Ono Y, Pauly M, Faber C, Glaesener R, de Goeij AF, Arends JW, Andersen SN, Lövig T, Breivik J, Gaudernack G, Clausen OP, De Angelis PD, Meling GI, Rognum TO, Smith R, Goh HS, Font A, Rosell R, Sun XF, Zhang H, Benhattar J, Losi L, Lee JQ, Wang ST, Clarke PA, Bell S, Quirke P, Bubb VJ, Piris J, Cruickshank NR, Morton D, Fox JC, Al-Mulla F, Lees N, Hall CN, Snary D, Wilkinson K, Dillon D, Costa J, Pricolo VE, Finkelstein SD, Thebo JS, Senagore AJ, Halter SA, Wadler S, Malik S, Krtolica K, Urošević N (2001) Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study. *Br J Cancer* **85**: 692–696
- Barber TD, Vogelstein B, Kinzler KW, Velculescu VE (2004) Somatic mutations of EGFR in colorectal cancers and glioblastomas. *N Engl J Med* **351**: 2883
- Baselga J, Arteaga CL (2005) Critical update and emerging trends in epidermal growth factor receptor targeting in cancer. *J Clin Oncol* **23**: 2445–2459
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc B* **57**: 289–300
- Benvenuti S, Sartore-Bianchi A, Di Nicolantonio F, Zanon C, Moroni M, Veronese S, Siena S, Bardelli A (2007) Oncogenic activation of the RAS/RAF signaling pathway impairs the response of metastatic colorectal cancers to anti-epidermal growth factor receptor antibody therapies. *Cancer Res* **67**: 2643–2648
- Brookmeyer R, Crowley J (1982) A confidence interval for the median survival time. *Biometrics* **38**: 29–41
- Cappuzzo F, Finocchiaro G, Rossi E, Jänne PA, Carnaghi C, Calandri C, Bencardino K, Ligorio C, Ciardiello F, Pressiani T, Destro A, Roncalli M, Crino L, Franklin WA, Santoro A, Varella-Garcia M (2008) EGFR FISH assay predicts for response to cetuximab in chemotherapy refractory colorectal cancer patients. *Ann Oncol* **19**: 717–723
- Chung KY, Shia J, Kemeny NE, Shah M, Schwartz GK, Tse A, Hamilton A, Pan D, Schrag D, Schwartz L, Klimstra DS, Fridman D, Kelsen DP, Saltz LB (2005) Cetuximab shows activity in colorectal cancer patients with tumors that do not express the epidermal growth factor receptor by immunohistochemistry. *J Clin Oncol* **23**: 1803–1810
- Clark-Langone KM, Wu JY, Sangli C, Chen A, Snable JL, Nguyen A, Hackett JR, Baker J, Yothers G, Kim C, Cronin MT (2007) Biomarker discovery for colon cancer using a 761 gene RT-PCR assay. *BMC Genomics* **8**: 279
- Cunningham D, Humblet Y, Siena S, Khayat D, Bleiberg H, Santoro A, Bets D, Mueser M, Harstrick A, Verslype C, Chau I, Van Cutsem E (2004) Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N Engl J Med* **351**: 337–345
- De Roock W, Piessevaux H, De Schutter J, Janssens M, De Hertogh G, Personeni N, Biesmans B, Van Laethem J, Peeters M, Humblet Y, Van Cutsem E, Tejpar S (2008) KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. *Ann Oncol* **19**: 508–515
- Di Fiore F, Blanchard F, Charbonnier F, Le Pessot F, Lamy A, Galais MP, Bastit L, Killian A, Sesboué R, Tuech JJ, Queuniet AM, Paillot B, Sabourin JC, Michot F, Michel P, Frebourg T (2007) Clinical relevance of KRAS mutation detection in metastatic colorectal cancer treated by cetuximab plus chemotherapy. *Br J Cancer* **96**: 1166–1169
- Di Nicolantonio F, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P, De Dossò S, Mazzucchelli L, Frattini M, Siena S, Bardelli A (2008) Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* **26**: 5705–5712
- Freeman DJ, Juan T, Reiner M, Hecht JR, Meropol NJ, Berlin J, Mitchell E, Sarosi I, Radinsky R, Amado RG (2008) Association of K-ras mutational status and clinical outcomes in patients with metastatic colorectal cancer receiving panitumumab alone. *Clin Colorectal Cancer* **7**: 184–190
- Guzinska-Ustymowicz K, Pryczynicz A, Kemona A (2009) PTP4A3 expression increases strongly in lymph node metastases from colorectal carcinoma. *Anticancer Res* **29**: 3913–3916
- Jacobs B, De Roock W, Piessevaux H, Van Oirbeek R, Biesmans B, De Schutter J, Fieuws S, Vandesompele J, Peeters M, Van Laethem JL, Humblet Y, Pénault-Llorca F, De Hertogh G, Laurent-Puig P, Van Cutsem E, Tejpar S (2009) Amphiregulin and epiregulin mRNA expression in primary tumors predicts outcome in metastatic colorectal cancer treated with cetuximab. *J Clin Oncol* **27**: 5068–5074
- Jonker DJ, Karapetis C, Harbison C, O'Callaghan CJ, Tu D, Simes RJ, Xu L, Moore MJ, Zalberg JR, Khambata-Ford S (2009) High epiregulin (EREG) gene expression plus K-ras wild-type (WT) status as predictors of cetuximab benefit in the treatment of advanced colorectal cancer (ACRC): results from NCIC CTG.CO.17—A phase III trial of cetuximab versus best supportive care (BSC) 2009. *ASCO Ann Meeting Proc*. Chicago, IL (Abstract 4016)
- Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, Simes RJ, Chalchal H, Shapiro JD, Robitaille S, Price TJ, Shepherd L, Au H, Langer C, Moore MJ, Zalberg JR (2008) K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* **359**: 1757–1765
- Khambata-Ford S, Garrett CR, Meropol NJ, Basik M, Harbison CT, Wu S, Wong TW, Huang X, Takimoto CH, Godwin AK, Tan BR, Krishnamurthi SS, Burris HA, Poplin EA, Hidalgo M, Baselga J, Clark EA, Mauro DJ (2007) Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol* **25**: 3230–3237
- Lièvre A, Bacht J, Boige V, Cayre A, Le Corre D, Buc E, Ychou M, Bouché O, Landi B, Louvet C, André T, Bibeau F, Diebold M, Rougier P, Ducreux M, Tomasic G, Emile J, Penault-Llorca F, Laurent-Puig P (2008) KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *J Clin Oncol* **26**: 374–379
- Loupakis F, Ruzzo A, Cremolini C, Vincenzi B, Salvatore L, Santini D, Masi G, Stasi I, Canestrari E, Rulli E, Floriani I, Bencardino K, Galluccio N, Catalano V, Tonini G, Magnani M, Fontanini G, Basolo F, Falcone A, Graziano F (2009) KRAS codon 61, 146 and BRAF mutations predict resistance to cetuximab plus irinotecan in KRAS codon 12 and 13 wild-type metastatic colorectal cancer. *Br J Cancer* **101**: 715–721
- Mendelsohn J, Baselga J (2006) Epidermal growth factor receptor targeting in cancer. *Semin Oncol* **33**: 369–385
- Meropol NJ (2005) Epidermal growth factor receptor inhibitors in colorectal cancer: it's time to get back on target. *J Clin Oncol* **23**: 1791–1793
- Modjtahedi H, Essapen S (2009) Epidermal growth factor receptor inhibitors in cancer treatment: advances, challenges and opportunities. *Anticancer Drugs* **20**: 851–855
- Molleví DG, Aytes A, Padullés L, Martínez-Iniesta M, Baixeras N, Salazar R, Ramos E, Figueras J, Capella G, Villanueva A (2008) PRL-3 is essentially overexpressed in primary colorectal tumours and associates with tumour aggressiveness. *Br J Cancer* **99**: 1718–1725
- Morlan J, Baker J, Sinicropi D (2009) Mutation detection by real-time PCR: a simple, robust and highly selective method. *PLoS ONE* **4**: e4584
- Moroni M, Veronese S, Benvenuti S, Marrapese G, Sartore-Bianchi A, Di Nicolantonio F, Gambacorta M, Siena S, Bardelli A (2005) Gene copy number for epidermal growth factor receptor (EGFR) and clinical

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

- response to antiEGFR treatment in colorectal cancer: a cohort study. *Lancet Oncol* **6**: 279–286
- Perrone F, Lampis A, Orsenigo M, Di Bartolomeo M, Gevorgyan A, Losa M, Frattini M, Riva C, Andreola S, Bajetta E, Bertario L, Leo E, Pierotti MA, Pilotti S (2009) PI3KCA/PTEN deregulation contributes to impaired responses to cetuximab in metastatic colorectal cancer patients. *Ann Oncol* **20**: 84–90
- Personeni N, Fieuws S, Piessevaux H, De Hertogh G, De Schutter J, Biesmans B, De Roock W, Capoen A, Debiec-Rychter M, Van Laethem J, Peeters M, Humblet Y, Van Cutsem E, Tejpar S (2008) Clinical usefulness of EGFR gene copy number as a predictive marker in colorectal cancer patients treated with cetuximab: a fluorescent *in situ* hybridization study. *Clin Cancer Res* **14**: 5869–5876
- Prenen H, De Schutter J, Jacobs B, De Roock W, Biesmans B, Claes B, Lambrechts D, Van Cutsem E, Tejpar S (2009) PIK3CA mutations are not a major determinant of resistance to the epidermal growth factor receptor inhibitor cetuximab in metastatic colorectal cancer. *Clin Cancer Res* **15**: 3184–3188
- Roth AD, Tejpar S, Delorenzi M, Yan P, Fiocca R, Klingbiel D, Dietrich D, Biesmans B, Bodoky G, Barone C, Aranda E, Norlinger B, Cisar L, Labianca R, Cunningham D, Van Cutsem E, Bosman F (2010) Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J Clin Oncol* **38**: 466–474
- Saltz LB, Meropol NJ, Loehrer PJ, Needle MN, Kopit J, Mayer RJ (2004) Phase II trial of cetuximab in patients with refractory colorectal cancer that expresses the epidermal growth factor receptor. *J Clin Oncol* **22**: 1201–1208
- Samowitz WS, Curtin K, Schaffer D, Robertson M, Leppert M, Slattery ML (2000) Relationship of Ki-ras mutations in colon cancers to tumor location, stage, and survival: a population-based study. *Cancer Epidemiol Biomarkers Prev* **9**: 1193–1197
- Samowitz WS, Sweeney C, Herrick J, Albertsen H, Levin TR, Murtaugh MA, Wolff RK, Slattery ML (2005) Poor survival associated with the BRAF V600E mutation in microsatellite-stable colon cancers. *Cancer Res* **65**: 6063–6069
- Sartore-Bianchi A, Martini M, Molinari F, Veronese S, Nichelatti M, Artale S, Di Nicolantonio F, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S, Bardelli A (2009) PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. *Cancer Res* **69**: 1851–1857
- Souglakos J, Philips J, Wang R, Marwah S, Silver M, Tzardi M, Silver J, Ogino S, Hooshmand S, Kwak E, Freed E, Meyerhardt JA, Saridaki Z, Georgoulas V, Finkelstein D, Fuchs CS, Kulke MH, Shivdasani RA (2009) Prognostic and predictive value of common mutations for treatment response and survival in patients with metastatic colorectal cancer. *Br J Cancer* **101**: 465–472
- Tsuchihashi Z, Khambata-Ford S, Hanna N, Jänne PA (2005) Responsiveness to cetuximab without mutations in EGFR. *N Engl J Medicine* **353**: 208–209
- Van Cutsem E, Dico M, Arber N, Berlin J, Cervantes A, Ciardiello F, De Gramont A, Diaz-Rubio E, Ducreux M, Geva R, Glimelius B, Glynne Jones R, Grothey A, Gruenberger T, Haller D, Haustermans K, Labianca R, Lenz HJ, Minsky B, Nordlinger B, Ohtsu A, Pavlidis N, Rougier P, Schmiegel W, Van de Velde C, Schmoll HJ, Sobrero A, Tabernero J (2010) Molecular markers and biological targeted therapies in metastatic colorectal cancer: expert opinion and recommendations derived from the 11th ESMO/World Congress on Gastrointestinal Cancer, Barcelona, 2009. *Ann Oncol* (Suppl 6): vi1–vi10
- Weinstein IB (2002) Cancer. Addiction to oncogenes—the Achilles heel of cancer. *Science* **297**: 63–64



This work is licensed under the Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>