OPEN

Oncogene (2016) **35**, 6403–6415

www.nature.com/onc

ORIGINAL ARTICLE

Integrated genomic analysis of colorectal cancer progression reveals activation of EGFR through demethylation of the *EREG* promoter

X Qu[1,8], T Sandmann[2,8], H Frierson Jr[3,8], L Fu[1], E Fuentes[4], K Walter[1], K Okrah[5], C Rumpel[3], C Moskaluk[3], S Lu[1], Y Wang[1], R Bourgon[2], E Penuel[1], A Pirzkall[6], L Amler[1], MR Lackner[1], J Tabernero[7], GM Hampton[1] and O Kabbarah[1]

Key molecular drivers that underlie transformation of colonic epithelium into colorectal adenocarcinoma (CRC) are well described. However, the mechanisms through which clinically targeted pathways are activated during CRC progression have yet to be elucidated. Here, we used an integrative genomics approach to examine CRC progression. We used laser capture microdissection to isolate colonic crypt cells, differentiated surface epithelium, adenomas, carcinomas and metastases, and used gene expression profiling to identify pathways that were differentially expressed between the different cell types. We identified a number of potentially important transcriptional changes in developmental and oncogenic pathways, and noted a marked upregulation of *EREG* in primary and metastatic cancer cells. We confirmed this pattern of gene expression by *in situ* hybridization and observed staining consistent with autocrine expression in the tumor cells. Upregulation of *EREG* during the adenoma–carcinoma transition was associated with demethylation of two key sites within its promoter, and this was accompanied by an increase in the levels of epidermal growth factor receptor (EGFR) phosphorylation, as assessed by reverse-phase protein analysis. In CRC cell lines, we demonstrated that *EREG* demethylation led to its transcriptional upregulation, higher levels of EGFR phosphorylation, and sensitization to EGFR inhibitors. Low levels of *EREG* methylation in patients who received cetuximab as part of a phase II study were associated with high expression of the ligand and a favorable response to therapy. Conversely, high levels of promoter methylation and low levels of *EREG* expression were observed in tumors that progressed after treatment. We also noted an inverse correlation between *EREG* methylation and expression levels in several other cancers, including those of the head and neck, lung and bladder. Therefore, we propose that upregulation of *EREG* expression through promoter demethylation might be an important means of activating the EGFR pathway during the genesis of CRC and potentially other cancers.

Oncogene (2016) **35,** 6403–6415; doi:10.1038/onc.2016.170; published online 6 June 2016

INTRODUCTION

The development of colorectal cancer (CRC) is known to proceed through the acquisition of genetic alterations during disease progression.[1] In colonic adenomas, there is disruption of the function of tumor suppressor gene, *APC*, as well as activating mutations in oncogenes, such as *KRAS*. Later events during the transition to carcinoma include loss of tumor suppressor gene, *TP53*.[1] As significant progress has been made towards understanding the biology of normal colonic epithelium,[2,3] it has become evident that many of the pathways regulating normal colonic surface and crypt homeostasis are also involved in oncogenic transformation, including the WNT, NOTCH, TGF-β (transforming growth factor-β), MAPK (mitogen-activated protein kinase) and PI3K/AKT (phosphatidylinositol 3-kinase/AKT) pathways.[1–10] Genomic aberrations, including mutations, microsatellite instability (MSI) and chromosomal instability (CIN) are known drivers of colonic epithelial transformation, and DNA methylation also contributes to disease development.[2,4,11–14] This has led to two molecularly defined subsets of CRC being described in recent years. One includes the CpG island methylator phenotype (CIMP) subtype that often exhibits microsatellite instability (MSI), a high frequency of *BRAF* mutation and represents ~15% of CRC.[14] The other subset is defined by CIN/*TP53* that frequently carries *KRAS* mutations and accounts for ~85% of CRCs.[14]

While CIMP and CIN/*TP53* subtypes encompass molecular events of significance in CRC, activation of receptor tyrosine kinase signaling has also been shown to have an important role in driving colon carcinogenesis and associated angiogenesis.[6,14,15] Indeed, the two classes of clinically approved therapies in CRC are antagonists of the vascular endothelial growth factor/receptor-2 (VEGF/VEGFR2) and epidermal growth factor receptor (EGFR) receptor tyrosine kinase signaling pathways, both of which are typically used in combination with fluorouracil-containing chemotherapy.[16–18] Patients with *RAS*/*RAF* mutant tumors do not usually respond well to EGFR-targeted therapies but do experience clinical benefit when treated with antiangiogenic drugs, such

[1]Oncology Biomarker Development, Genentech Inc., South San Francisco, CA, USA; [2]Bioinformatics and Computational Biology, Genentech Inc., South San Francisco, CA, USA; [3]Department of Pathology, University of Virginia Health System, Charlottesville, VA, USA; [4]Department of Pathology, Genentech Inc., South San Francisco, CA, USA; [5]Department of Biostatistics, Genentech Inc., South San Francisco, CA, USA; [6]Department of Clinical Sciences, Genentech Inc., South San Francisco, CA, USA and [7]Department of Medical Oncology, Vall d'Hebron University Hospital, Barcelona, Spain. Correspondence: Dr GM Hampton or Dr O Kabbarah, Oncology Biomarker Development, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Email: garreth@gene.com (GMH) or kabbarao@gene.com (OK)
[8]These authors contributed equally to this work.
Received 20 October 2015; revised 26 February 2016; accepted 12 April 2016; published online 6 June 2016

as avastin.[19–22] Conversely, patients with *RAS*/*RAF* wild-type tumors have been shown to respond favorably to EGFR antagonistic antibodies, such as cetuximab.[19–21,23] Retrospective analyses have also suggested that patients with *RAS* wild-type tumors that express high level of the EGFR ligands, EREG and AREG, might benefit from cetuximab treatment.[21,24,25] However, the timing and mechanism through which the EGFR pathway is activated during CRC progression have yet to be revealed.

In this study, we examined CRC progression using an integrative genomic approach. We observed broad transcriptional differences between laser capture-microdissected (LCM) normal colonic surface epithelium, crypt cells, adenomas and CRCs in pathways known to be involved in cell proliferation, differentiation and transformation. Here, we focused on the clinically relevant EGFR pathway because of the marked upregulation of the gene encoding for the EGFR ligand, EREG, that we observed at the adenoma–carcinoma transition. Mechanistically, we found *EREG*-mediated activation of the EGFR pathway in CRC to be associated with demethylation of its promoter. We demonstrated in CRC cell lines that global demethylation released epigenetic inhibition of *EREG* and led to higher levels of EGFR phosphorylation, as well as increased sensitization to EGFR inhibitors. In patients who received cetuximab as part of a phase II trial, we observed low levels of *EREG* methylation and high level of ligand expression in tumors that exhibited the best responses. Finally, we detected an inverse correlation between *EREG* methylation and expression levels in different tumor types, suggesting that epigenetic regulation of *EREG* expression might be a common mechanism for EGFR pathway activation in several types of malignancies.

RESULTS

An integrative molecular view of colorectal cancer progression

To gain a molecular understanding of normal colonic epithelial biology and CRC progression, we used an integrative genomics approach. First, we used LCM to isolate cells from normal colonic crypts ($n=7$), normal colonic surface epithelium ($n=6$), colonic adenomas ($n=17$), primary colorectal carcinomas ($n=17$) and CRC distant metastases ($n=11$) (Figure 1a and Supplementary Table S1). To obtain a broad view of the molecular patterns from normal and pathophysiological cell types, we performed gene expression profiling analysis. Because of the limited amount of material available from these samples, we performed targeted next-generation sequencing, genome-wide methylation and reverse-phase protein array (RPPA) analyses on a series of 48 macrodissected frozen samples of normal colonic epithelium ($n=14$), adenomas ($n=12$), primary colorectal

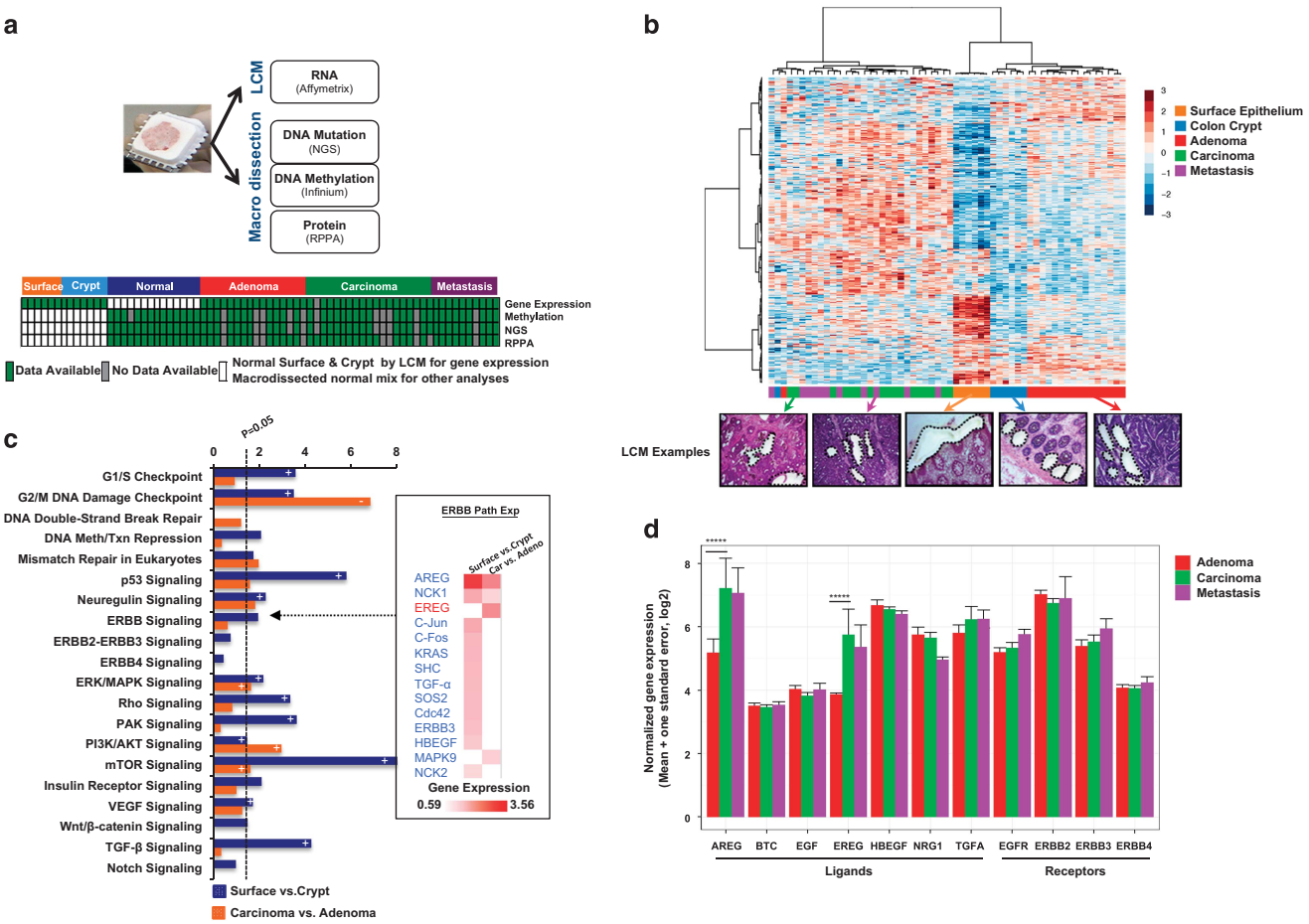

Figure 1. Integrative analysis of colorectal cancer progression. (**a**) Summary of sample characteristics and assays run on each sample. (**b**) Hierarchical clustering of the gene expression scores for the top 500 most variable genes across all samples (after adjustment for one surrogate variable). Below heatmap: Examples of LCM cells from normal surface epithelium, colon crypts, adenomas, carcinomas and metastasis within black dotted areas. (**c**) Selected differentially expressed signaling pathways between normal surface and normal crypt, and/or between adenoma and carcinoma, and differentially expressed genes annotated as ERBB signaling components by Ingenuity analysis ($P<0.05$, log 2 fold change >0.58). Predicted direction of pathway activation based on Ingenuity analysis is reflected by a +/− sign on the respective bars. (**d**) Bar plots representing expression levels of ERBB family ligands and receptors in adenomas, carcinomas and metastases. Two-sided *P*-values were derived using an unpaired *t*-test. *****$P<0.00005$.

carcinomas ($n = 13$) and CRC metastases ($n = 9$) (Figure 1a and Supplementary Table S2).

To establish a transcriptional view of normal and malignant colonic development, we used Affymetrix microarrays and profiled LCM tissue samples. Unsupervised hierarchical clustering based on the expression of 500 genes that exhibited the most variation in expression across all samples revealed a number of notable features. Overall, samples segregated into two major groupings: (1) carcinomas and metastases interspersed with one another, and (2) normal crypts cells, surface epithelial cells and adenomas (Figure 1b). Although crypt and surface epithelial cells cosegregated, they showed some of the most striking differences in gene expression, as clearly demonstrated by principal component analysis (Figure 1b and Supplementary Figure S1), reflecting undifferentiated cells within the crypt environment and differentiated surface epithelial. That carcinomas and metastases showed a lack of significant differences in gene expression (as judged by their cosegregation) (Figure 1b) is consistent with the concept that the time from carcinoma and metastasis is significantly shorter (< 2 years) than the time interval from adenoma to carcinoma (~17 years),[26] and is suggestive of few additional alterations occurring during metastatic seeding and growth.

To understand the key molecular features underlying the differences in cell types, we used the 1416 differentially expressed genes between normal surface epithelium and crypt cells, and 643 differentially expressed genes between adenomas and carcinomas at a false discovery rate of $< 5\%$ and absolute log 2 fold change > 0.58 (Supplementary Table S3), and queried them in the canonical pathway annotation in Ingenuity analysis. Several pathways known to be involved in colonic cell differentiation were differentially expressed between surface epithelium and the crypt compartments, including the WNT/β-catenin, TGF-β and G1/S cell cycle checkpoint, as well as the mTOR, TP53 and DNA methylation and double-stranded break repair pathways[2,3,5,7–10,12,27,28] (Figure 1c and Supplementary Table S4). At the adenoma–carcinoma transition, genes representative of the G2/M checkpoint and PI3K/AKT pathways[29] were more prominently differentially expressed (Figure 1c). Interestingly, genes belonging to the ERBB pathway exhibited a mixed pattern of expression that indicated preferential activation of the ERBB2 pathway in surface epithelium compared with crypt cells, and a similar degree of neuregulin pathway activation in surface/crypt and adenomas/carcinomas (Figure 1c), suggesting that ERBB pathway may have a role in both normal colon epithelial cell differentiation and malignant transformation.

We used a custom next-generation sequencing cancer panel[30] to assess the mutational status of CRC-relevant genes in our progression sample set, and detected mutations in *APC*, *KRAS*, *PIK3CA, PTEN*, *SMAD4* and *TP53* (Supplementary Figure S2). The temporal occurrence of mutations was consistent with the reported timing of these genetic alterations during CRC progression.[1] For example, we noted the presence of *APC* and *KRAS* mutations in adenomas, whereas *TP53* mutations were detected in carcinomas (Supplementary Figure S2). Thus, our targeted next-generation sequencing data recapitulates the presence and timing of previously described mutations, and suggests that our cohort is suitable for discovery of molecular alteration associated with the genesis of CRC.

As depicted in Figure 1c, we identified multiple pathways that showed differential expression between the different tissue types; however, given the therapeutic importance of EGFR in CRC, we examined the expression of members of this pathway in our LCM tissues. Several genes encoding EGFR pathway components were differentially represented between surface epithelium and crypt cells, including the ligands *AREG* and *TGF-α*, the *ERBB3* receptor, as well as downstream components such as *SHC*, *NCK1*, *NCK2*, *c-Fos* and *c-Jun* (Figure 1c). Of these genes, only *AREG* and *NCK1* exhibited differential expression between adenoma and carcinomas, and *EREG* was upregulated specifically in carcinomas (Figure 1c). A closer examination of the expression levels of ERBB family ligands and receptors[31] in adenomas compared with carcinomas and metastases showed no significant increase in any of the four receptors and most of the ligands, and a marginal, but not significant, increase in the levels of TGF-α in adenomas compared with carcinomas and metastases (Figure 1d). On the other hand, *AREG* and *EREG* were expressed at significantly higher levels in carcinomas versus adenomas (Figures 1c and d and Supplementary Table S3). Interestingly, *EREG* was upregulated in carcinomas compared with adenomas but was not differentially expressed between colonic surface epithelium and crypt cells (Figures 1c and d and Supplementary Table S3), pointing to the possibility that AREG and EREG might be having different roles in normal colonic epithelial cell differentiation and in CRC development.

EREG transcript is detected at low levels in non-malignant cells and its levels are markedly increased in carcinomas

Because we specifically isolated adenomatous and tumor cells by LCM, we hypothesized that expression of EGFR ligands likely occurred in a cell-autonomous manner. To test this hypothesis, we performed *in situ* hybridization (ISH) analysis using a custom assay on tissues from an independent-sample set (Supplementary Table S5). EREG transcript number and localization were quantified in matched normal, adenoma and carcinoma tissues from the same patients (see Materials and methods). A representative image showing EREG transcript signals in a case that included normal colonic epithelium, adenoma and carcinoma histologies in the same tissue section is shown in Figure 2a. We found that normal colonic epithelial cells and adenomas showed low levels of EREG (Figure 2a). In contrast, a neighboring carcinoma exhibited significantly higher levels of EREG specifically in tumor cells (Figure 2a). Stromal cells showed minimal EREG ISH signal, irrespective of whether they were adjacent to normal epithelium, adenomas or carcinomas (Figures 2a and b). In the remaining samples, we observed that the EREG signal was low to absent in most normal colonic epithelia, stromal cells and adenomas, but was markedly higher in a majority of carcinomas (Figure 2b).

We also examined the levels and localization of the AREG transcript in the same cohort using a custom ISH assay. We found that AREG signals were significantly higher in carcinomas compared with adenomas (Supplementary Figure S3). However, unlike in the case of EREG, where expression was high in tumor cells and low to absent in adenomas and normal colonic epithelium, a wide range of AREG ISH signal was detected in normal colonic epithelia and in adenomas that was, in some cases, comparable to the levels typically observed in carcinomas (Supplementary Figure S3). AREG signal was not detected in stromal cells (Supplementary Figure S3). The observation that *EREG* was markedly upregulated and expressed specifically in tumor cells is consistent with the EGFR ligand as a potential driver of CRC development, and supports an autocrine model of pathway activation that has been proposed in head and neck and lung cancers.[32,33]

EREG promoter demethylation and transcriptional upregulation are associated with increased EGFR phosphorylation during the adenoma–carcinoma transition

Epigenetic regulation of gene expression has been shown to have an important role in CRC development.[34] To gain a broad view of genes that might be epigenetically regulated during CRC development, we examined the promoter methylation status within the promoter region 2 kb downstream to 0.5 kb upstream of transcriptional start sites in macrodissected tissues. In total,

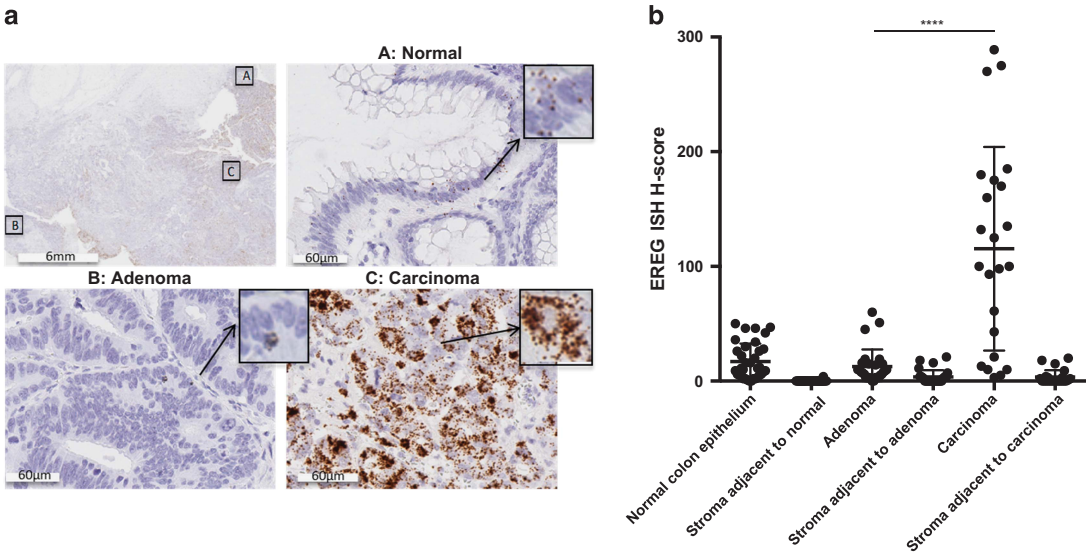

Figure 2. *EREG* expression is confined to colon epithelial cells and is markedly upregulated in carcinomas. (**a**) Representative ISH images of EREG transcript showing signal in normal colonic epithelium, adenoma, stromal cells and carcinomas. (**b**) Summary of EREG ISH scores in different cell types. Two-sided *P*-values were derived using an unpaired *t*-test, $n = 35$, 35, 27, 27, 23 and 23, respectively. ****$P < 0.0001$.

gene expression and promoter methylation data was available for 11 813 genes in 34 samples. We calculated Spearman's rank correlation between gene expression and methylation data across matched adenomas, carcinomas and metastases to identify genes that might be regulated by promoter methylation (Supplementary Table S6). Interestingly, the MAPK pathway activation suppressor, *dual specificity phosphatase 4*, was the top-scoring gene, exhibiting the highest negative correlation between promoter methylation and gene expression (Figure 3a). Other genes that exhibited the strongest inverse correlations between their levels of methylation and expression included *Disabled-2*, which is often silenced in gastrointestinal tumors,[35] and *S100 calcium binding protein A4*, which has been shown to interact with and degrade TP53[36] (Figure 3a). We found *EREG* to be one of the top-ranked genes that exhibited a large negative correlation coefficient (Figure 3a and Supplementary Table S6, 28 out of 11 813 genes), suggesting that the upregulation of *EREG* expression during CRC progression might be mediated by promoter demethylation.

To determine the sites responsible for epigenetic regulation of *EREG* expression in CRC, we examined the methylation status for all eight available probes within the promoter region. Among these two probes, cg11646192 and cg19308222, spaced 224bp apart, showed a strong inverse correlation between *EREG* expression and promoter methylation, and were highly correlated with each other (Figures 3b and c). The six remaining probes were consistently either hyper- or hypomethylated in all samples irrespective of the tissue type and levels of *EREG* expression (Figure 3b). We examined the methylation signals of the cg11646192 and cg19308222 probes in our CRC progression samples and found low levels of methylation in normal colon surface epithelium and crypt tissue, a high degree of methylation in adenomas and demethylation in primary tumors and metastases (Supplementary Figure S4). These data suggest that increased *EREG* expression during CRC development might be due to promoter demethylation at these two positions. Although the expression of both *EREG* and *AREG* was higher in carcinomas and metastases compared with that in adenomas, the methylation status of two *AREG* methylation probes did not show a significant correlation with its expression levels (Supplementary Figure S5).

We next asked if decreased methylation and increased *EREG* expression had an impact on the phosphorylation status of two key sites on the EGFR receptor (EGFR-Y1068 and -Y1173). These two sites have been shown to be phosphorylated and activated by EGFR ligands.[37–39] Using RRPA, we observed significantly higher levels of EGFR phosphorylation in carcinomas and metastases compared with adenomas at both sites (Figure 3d), consistent with EGFR signaling being activated at the adenoma–carcinoma transition and maintained in distant CRC metastases. Moreover, a significant inverse correlation between *EREG* promoter methylation and EGFR phosphorylation was also observed at both sites in adenomas, carcinomas and metastases (Figure 3e). Taken together, our data suggest that activation of EGFR signaling is highly correlated with demethylation of specific sites within the *EREG* promoter and concomitant upregulation of the ligand during CRC progression.

Validation of the inverse correlation between *EREG* methylation and expression during CRC development

To validate the relationship between *EREG* demethylation and expression during CRC progression, we used an independent set of samples from 16 patients with matched adenomas and carcinomas (Supplementary Table S5). We measured the percentage of *EREG* promoter methylation using a custom methylation-specific PCR assay,[40] and assessed *EREG* expression levels using a commercial TaqMan Real-Time PCR assay (see Materials and methods). Consistent with our observations in the discovery cohort, *EREG* methylation was higher in adenomas compared with matching carcinomas, with concomitantly higher ligand expression levels in a subset of carcinomas compared with that in matched adenomas (Figure 4a). Notably, *EREG* expression was highest in carcinomas with the lowest levels of promoter methylation (Figure 4a). In aggregate, we detected significantly higher levels of *EREG* promoter methylation in adenomas compared with carcinomas (Figure 4b), and significantly higher transcript levels in carcinomas versus adenomas (Figure 4c). These results independently confirm that a decreased state of *EREG* methylation is often associated with increased levels of ligand expression during the adenoma–carcinoma transition.

Given that CIMP defines a distinct molecular CRC subtype that is known to be associated with response to 5-fluorouracil-based therapies,[14,41,42] we next examined the relationship between CIMP status and *EREG* methylation in CRC tissues from the Cancer

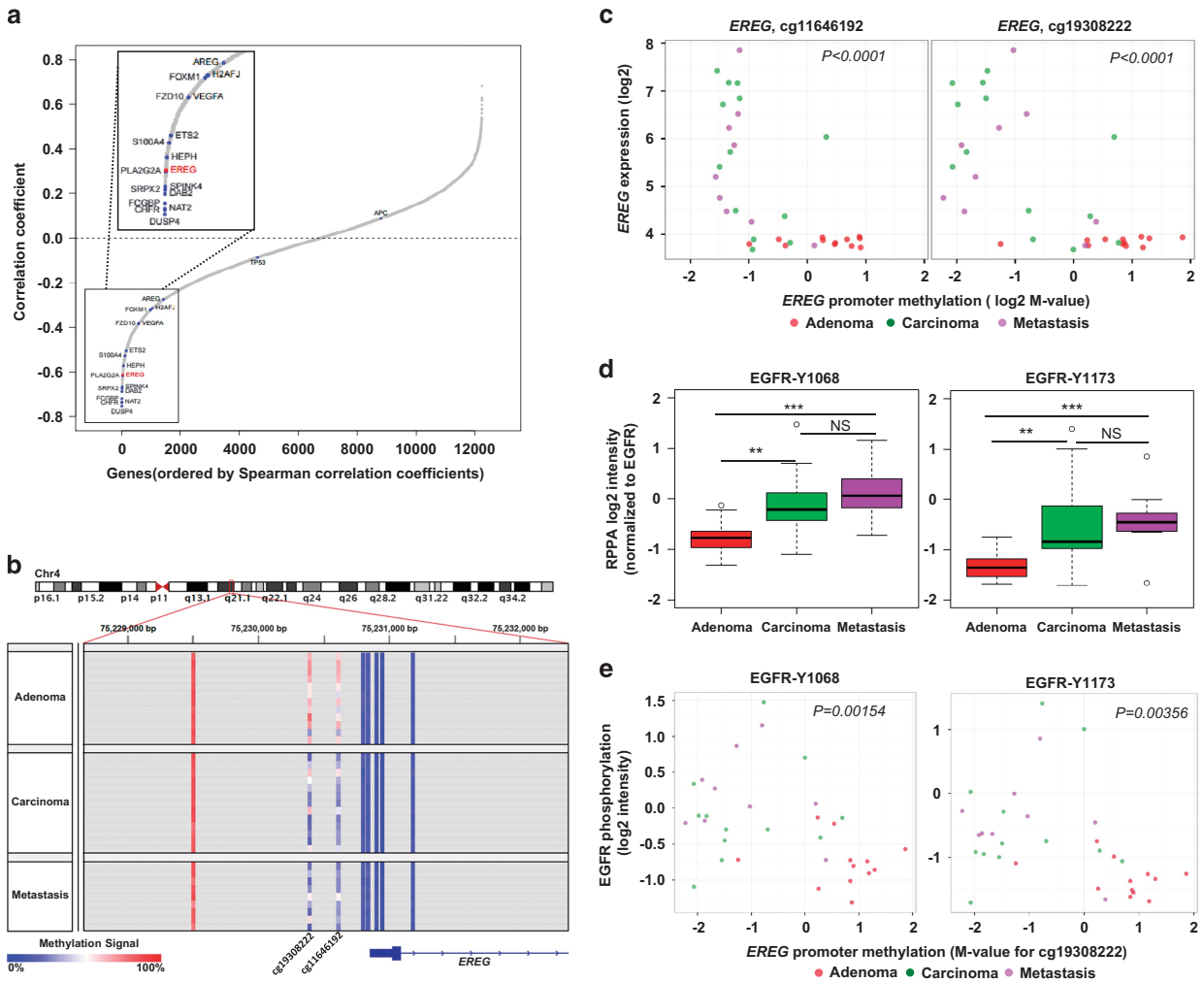

Figure 3. Activation of the EGFR pathway through *EREG* demethylation during CRC progression. (**a**) Spearman's rank correlation between normalized gene expression and mean promoter probe methylation (mean M-value) across all genes and all samples assayed on both Affymetrix gene expression and Illumina Infinium methylation microarrays. (**b**) *EREG* methylation within its promoter region (2 kb downstream–0.5 kb upstream of transcriptional start) across samples. (**c**) Correlation between *EREG* expression and promoter methylation at probes cg.11646192 and cg.19308222 in adenomas, carcinomas and metastases. Spearman's correlations $R = -0.68$ and -0.63, respectively. (**d**) Reverse-phase protein array analysis of phosphorylated EGFR (normalized to total EGFR) during CRC progression. (**e**) Correlation between *EREG* promoter methylation at probe cg.19308222 and EGFR phosphorylation at two key sites EGFR-Y1068 and -Y1173 in adenomas, carcinomas and metastases. Spearman's correlations $R = -0.53$ and -0.50, respectively. Two-sided *P*-values were derived using an unpaired *t*-test, $n = 33$. $^{**}P < 0.005$, $^{***}P < 0.0005$ and NS, not significant.

Genome Atlas Network (TCGA) collection. Because CIMP status was not available for all samples and given the high degree of concordance between CIMP status and *MLH1* promoter methylation that has been described previously,[14] we used *MLH1* promoter methylation as a surrogate for CIMP status (see Materials and methods). As expected, *EREG* methylation levels were significantly higher in CIMP+ compared with CIMP− samples (Supplementary Figure S6a). Consistently, *EREG* expression levels were significantly lower in CIMP+ compared with CIMP− samples (Supplementary Figure S6b). A similar relationship between CIMP status and *AREG* expression was also observed (Supplementary Figures S6c and d). Interestingly, ~25% of CIMP− samples exhibited high levels of methylation and low levels of *EREG* expression (Supplementary Figures S6a and b), suggesting that epigenetic regulation of *EREG* can occur in CIMP− samples and that epigenetic control of *EREG* cannot be explained solely by CIMP.

Our validation studies in an independent set of matched adenomas and carcinomas and from TCGA tissues are consistent with our finding that *EREG* expression can be epigenetically regulated during CRC progression, and suggest that this phenomenon can occur in both CIMP+ and CIMP− disease.

Demethylation and transcriptional upregulation of *EREG* leads to EGFR activation and sensitizes CRC cell lines to EGFR inhibitors

To begin to assess the role of methylation in regulating *EREG* expression and EGFR signaling, we next examined the expression and methylation status of *EREG* in a panel of CRC cell lines. Consistent with what we observed in the discovery and independent validation tissues sets, we noted an inverse correlation between *EREG* expression and methylation cross 40 CRC cell lines (Supplementary Figure S7a and Supplementary Table S7). From these 40 cell lines, we selected a representative panel of 10 CRC cell lines that exhibited low, intermediate and high basal levels of *EREG* expression and, accordingly, high, intermediate and low basal levels of *EREG* methylation (Figure 5a). We next assessed the basal EGFR phosphorylation levels in these 10 cell lines at the EGFR-Y1068 and -Y1173 sites that are known to be indicative of receptor activation.[37–39] We found cell lines with

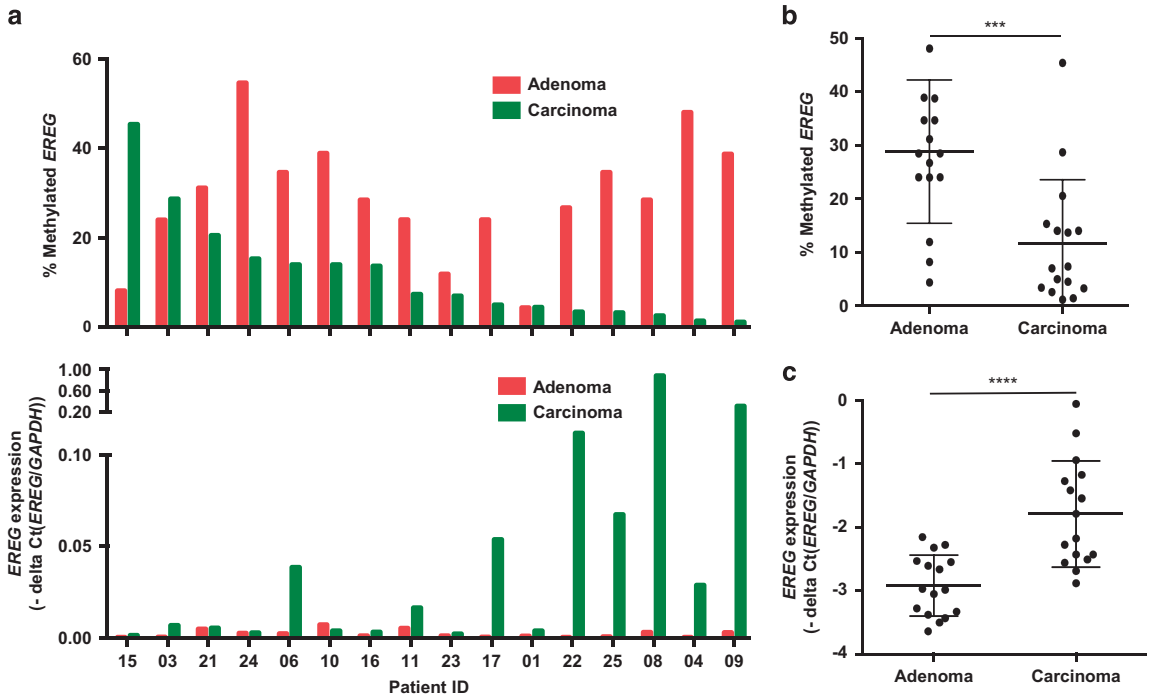

Figure 4. *EREG* expression and methylation in an independent patient set with matched adenoma and carcinoma tissues. (**a**) Percentage of *EREG* methylation using a custom assay, and relative *EREG* expression using a commercial Taqman assay in a series of matched adenomas and carcinomas. (**b**) Significantly higher levels of *EREG* methylation were observed in adenomas compared with carcinomas. (**c**) Significantly lower levels of *EREG* expression were detected in adenomas compared with carcinomas. Two-sided *P*-values were derived using a paired *t*-test, $n = 16$. ***$P < 0.001$ and ****$P < 0.0001$.

high basal levels of *EREG* methylation and low basal levels of expression (Colo302DM, Colo741 and RKO) to have barely detectable levels of EGFR phosphorylation at both sites (Figure 5c). In contrast, EREG phosphorylation was readily detectable in cell lines with intermediate levels of *EREG* expression and methylation (C2BBe1, CL11 and HCA7) and in cell lines with low basal levels of *EREG* methylation and intermediate/high *EREG* expression levels (HCT15, SW48, HCT116 and in DLD-1) (Figures 5a–c). These findings support a strong association between basal *EREG* methylation and expression status and EGFR phosphorylation levels in CRC cell lines.

To begin assessing the functional consequences of *EREG* methylation on EGFR phosphorylation in CRC cell lines, we treated the 10 representative CRC cell lines with the DNA demethylating agent, 5-aza-dC.[43] Following treatment, *EREG* methylation was significantly reduced in cells with high basal levels (Colo320DM, Colo741, RKO) and with intermediate basal levels (C2BBe1, CL11 and HCA7) of *EREG* methylation (Figure 5a). A significant increase in *EREG* expression levels was observed in all cell lines with the exception of HCA7, which exhibited a marginal increase of *EREG* expression (Figure 5b). In contrast, no induction of *EREG* expression was observed in cells lines with high basal levels of *EREG* expression (HCT15, SW48, HCT116 and DLD-1) (Figure 5b). After 5-aza-dC treatment, we also observed an increase in the levels of EGFR phosphorylation at both Y1068 and Y1173 sites in Colo741, RKO, C2BBe1 and CL11 cells (Figure 5c). No increase in EGFR tyrosine phosphorylation was observed in HCT15, SW48, HCT116 and DLD-1 cells (Figure 5c). These results suggest that DNA methylation most likely has an important role in regulating *EREG* expression and EGFR phosphorylation in CRC cell lines.

We next sought to determine whether the induction of *EREG* expression after 5-aza-dC treatment and the resulting increase in EGFR phosphorylation would enhance the sensitivity of CRC cell lines to the EGFR inhibitors, gefitinib and erlotinib. Unfortunately, very few CRC cell lines have been reported to be responsive to EGFR antagonists *in vitro*.[44] Among the four cell lines with high levels of endogenous *EREG* expression and EGFR phosphorylation, we chose SW48 as a control because it was the only sensitive line to treatment with EGFR inhibitors[44] (Figure 5d and Supplementary Figure S7b). We also selected cell line CL11 because it exhibited phospho-EGFR induction after 5-aza-dC treatment, which was comparable to the levels observed in SW48 cells. We pretreated SW48 and CL11 cells with 5-aza-dC for 2 days and then added 0.11, 0.33 or 1 μM of EGFR inhibitor alone or in combination with 5-aza-dC. Interestingly, a significant decrease of cell viability was observed in CL11 cells, which were treated with the combination of gefitinib and 5-aza–dC compared with those treated with gefitinib alone (Figure 5d). In contrast, no decrease in cell viability was observed in SW48 cells (Figure 5d). Similar results were observed in response to treatment with erlotinib (Supplementary Figure S7b). These data suggest that induction of *EREG* expression and EGFR phosphorylation after treatment of CL11 cells with 5-aza-dC sensitizes these cells to the EGFR inhibitors, gefitinib and erlotinib.

EREG methylation in CRC is associated with lack of clinical response to cetuximab treatment and may regulate *EREG* expression in several other types of cancers

EREG expression has been reported to predict clinical benefit in response to anti-EGFR treatment.[21,24,25,45] As we demonstrated that upregulation of *EREG* expression during CRC progression was associated with demethylation of its promoter in both our discovery and independent validation sets, we hypothesized that low methylation levels of *EREG* in tumors should predict favorable clinical response to cetuximab. To test this hypothesis, we analyzed primary tumor tissues from CRC patients with *KRAS* wild-type tumors who received cetuximab+FOLFIRI in the second-line metastatic setting as part of the MEHD7945A phase II clinical trial. We evaluated the methylation and expression levels of *EREG*

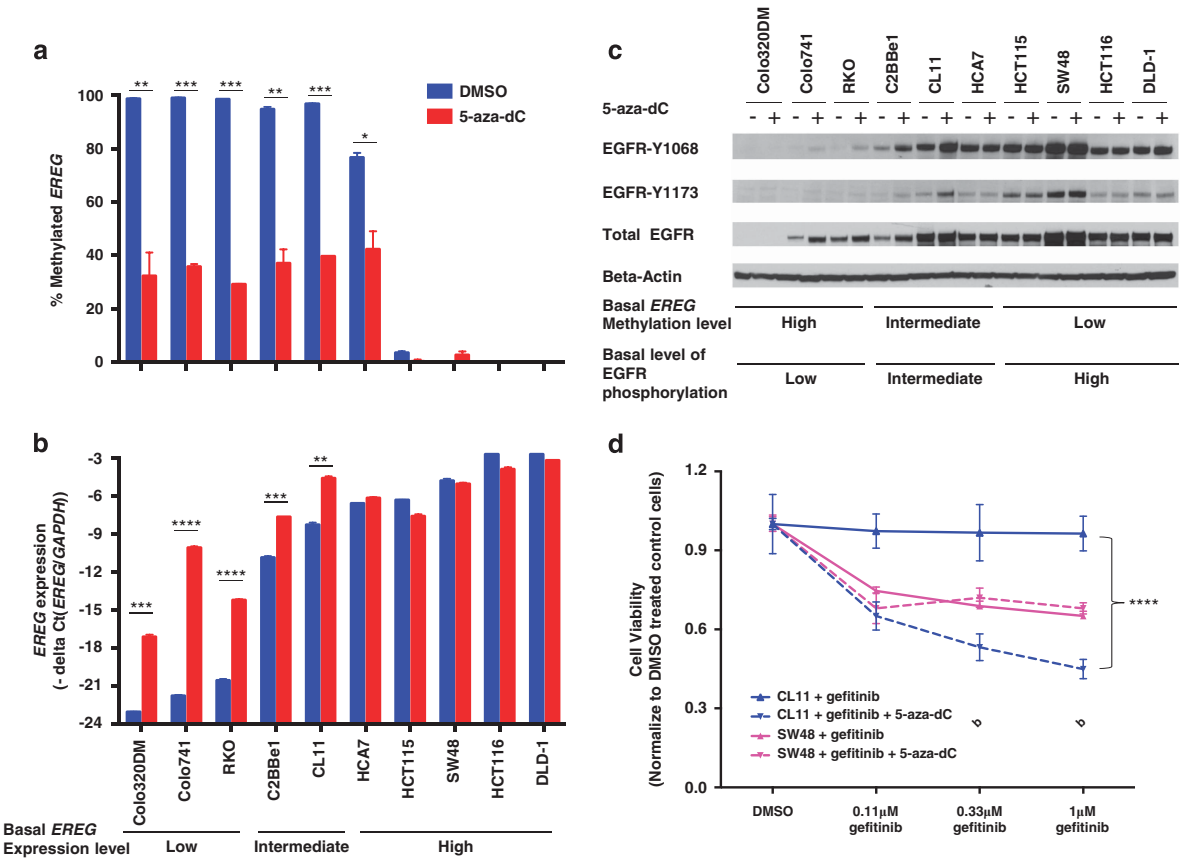

Figure 5. Demethylation and transcriptional upregulation of *EREG* leads to EGFR activation and sensitizes CRC cells to gefitinib. *EREG* expression is induced and the levels of EGFR phosphorylation are increased by 5-aza-dC treatment in colon cancer cell lines with low basal levels of EREG mRNA. (**a**) Percentage of *EREG* promoter methylation at probe cg.1930222, (**b**) *EREG* expression levels and (**c**) total EGFR and phospho-EGFR-Y1068 and -Y1173 levels in colon cancer cell lines after treatment with DMSO or 500nM 5-aza-dC for 4 days. Beta-Actin was used as a loading control. (**d**) 5-Aza-dC treatment significantly enhanced sensitivity to gefitinib in CL11 cells with high basal levels of methylation and low levels of *EREG* expression but had no impact on the growth of SW48 cells with low basal level of methylation and high levels of *EREG* expression. Two-sided *P*-values were derived using an unpaired *t*-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$. Results represent mean value ± s.e.m.

in tissues from 33 patients (Supplementary Table S8). Waterfall plots of best % helical computed tomography (CT) responses compared with baseline tumor measurements showed that six of seven (86%) of the patients who exhibited ⩾50% tumor size reduction had low levels of *EREG* promoter methylation (Figure 6a). Conversely, five of six (83%) of the patients whose tumors grew after treatment exhibited high levels of methylation (Figure 6a). Consistent with our earlier observations from the discovery and validation cohorts, *EREG* expression in samples from the phase II trial was inversely correlated with levels of promoter methylation (Figure 6b). Importantly, low levels of *EREG* promoter methylation were significantly correlated with tumor size reductions (Figure 6c). In line with these findings, increased *EREG* expression positively correlated with cetuximab activity (Figure 6c). Although *AREG* and *EREG* expression were significantly correlated in samples from the phase II study (Supplementary Figure S8a), AREG levels were not significantly associated with tumor size reduction following cetuximab+FOLFIRI treatment (Supplementary Figure S8b).

Increased levels of EREG have been reported in other cancer types;[46] however, the mechanism through which the ligand is regulated in these indications is largely unknown. To investigate whether methylation might be a mechanism for *EREG* regulation in cancers beyond CRC, we examined the relationship between *EREG* expression and methylation in samples from the TCGA data set.[45] We observed an inverse correlation between *EREG* expression and methylation in various cancer types, and, as expected, this was most notable in rectal and colon cancer (Figure 6d). Interestingly, head and neck cancers, lung adenocarcinomas, acute myeloid leukemias and squamous cell lung cancers all exhibited inverse correlations, indicating that *EREG* methylation might be regulating gene expression in these cancer types. Notably, these are indications in which anti-EGFR therapies have demonstrated antitumor activity.[40,45,47–49] In contrast, we observed a weaker inverse correlation between *AREG* methylation and expression across the TCGA indications (Supplementary Figure S9). Our data from the analysis of MEHD7945A samples and TCGA data sets suggest that *EREG* demethylation might be a mechanism for activating the EGFR pathway in CRC and, potentially, other types of cancers.

DISCUSSION

Our integrative molecular characterization of normal colonic surface epithelia, crypt cells, adenomas, primary carcinomas and distant metastases detected known and identified novel genetic drivers of CRC. Using targeted next-generation sequencing, we confirmed the presence of mutations in *APC* and *KRAS* genes as occurring early during the normal–adenoma transition, and *TP53* mutations as occurring late in CRC development.[1,26] On the transcriptional level, use of LCM enabled us to show that normal colonic surface epithelium and crypt cells are transcriptionally distinct from carcinomas and metastases, exhibiting wide-ranging

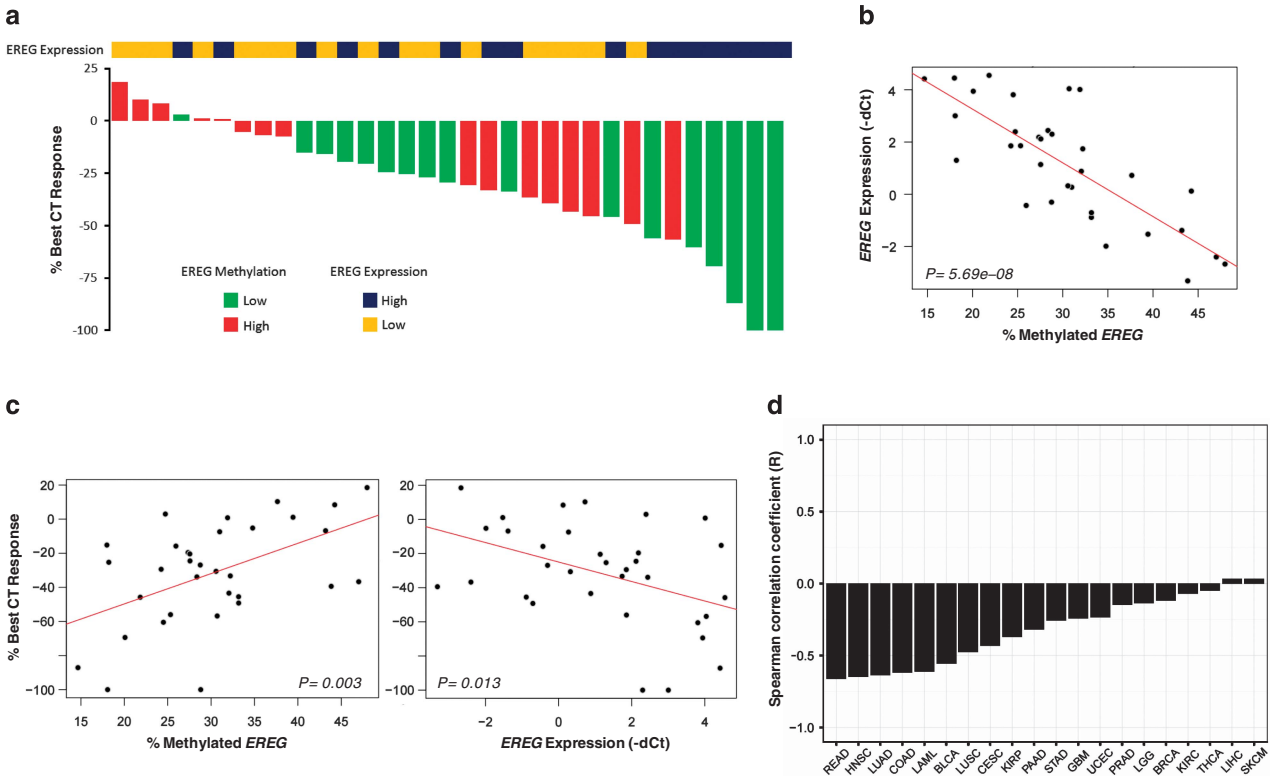

Figure 6. *EREG* methylation in CRC is associated with lack of response to cetuximab treatment and might be a mechanism for regulating *EREG* expression in several other cancer types. (**a**) Waterfall plot of % best CT response and associated *EREG* methylation and expression in samples from the cetuximab arm of an MEHD7945A+FOLFIRI phase II clinical trial. A median cutoff was used to designate samples as having high vs low % methylation or expression. (**b**) *EREG* methylation is inversely correlated with its expression in samples from the phase II clinical trial; $n = 33$ samples. Pearson's $R = -0.79$. (**c**) Association between *EREG* methylation and expression with % best CT response in the samples from the phase II clinical trial; $n = 33$ samples. Pearson's $R = 0.51$ and -0.43, respectively. (**d**) Correlation between *EREG* promoter methylation and expression across all TCGA indications with detectable EREG. AML, acute myeloid leukemia; BLCA, bladder urothelial carcinoma; BRCA, breast-invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, colon adenocarcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; PAAD, pancreatic adenocarcinoma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; THCA, thyroid carcinoma; UCEC, uterine corpus endometrioid carcinoma.

differences from them in the expression of genes belonging to developmental pathways including G1/S cell cycle checkpoint, TP53, WNT, TGF-β and mTOR. Hierarchical clustering indicated that gene expression in the normal crypt cells is more akin to that of adenomas than surface colonic epithelium, supporting the notion that adenomas and the normal crypt compartment might share a common origin, with clonal expansion of adenomas possibly resulting from mutations in stem cells within the crypts. We found that carcinomas and metastases were transcriptionally similar, consistent with previous findings that molecular determinants of metastasis are already present in primary carcinomas.[26]

Despite exhibiting many transcriptional differences, some pathways were consistently differentially expressed between both surface and crypt epithelium and also between adenomas and carcinomas, including EGFR-related neuregulin, and the ERK/MAPK pathways. Interestingly, a closer examination of EGFR pathway components revealed that *EREG* and *AREG* had similar expression profiles in CRC but different patterns of expression in normal colonic epithelium. Although *TGF-α* was upregulated during the adenoma–carcinoma transition, this change was not statistically significant. There are several reports that point to potentially distinct roles for TGF-α, AREG and EREG in ERBB pathway signaling in cancer.[46,50–52] In our data set, upregulation of *EREG* was observed in carcinomas compared with adenomas but not in surface epithelium compared with crypt epithelium, whereas *AREG* was consistently upregulated in both normal surface epithelium and in carcinomas. In line with our molecular data, examination of the spatial distribution of EREG by ISH showed minimal staining in normal surface and crypt epithelium, adenomas, and stromal cells, but strikingly high signals in carcinomas and metastases. Thus, the production of EREG by CRC cells appears to occur in an autocrine manner. Although *AREG* expression levels were significantly higher in tumors compared with other tissue types, AREG staining patterns were not tumor-specific, and the wide range in AREG signal intensity in normal colonic epithelium and adenomas was overlapping with that observed in carcinomas.

When we integrated genome-wide methylation and expression data to identify genes that might be epigenetically regulated, we found that *EREG* was one of the genes with the highest inverse correlation between methylation and expression levels. These findings were consistent with epigenetic silencing of *EREG* through methylation of the promoter in adenomas. *EREG* methylation was reversed during the transition to carcinoma, and this was associated with an upregulation of the ligand in cancers. In analyzing samples from our validation set of matched adenomas and carcinomas, we observed high levels of methylation and minimal expression of *EREG* in adenomas, whereas many matched carcinomas from the same patients exhibited low

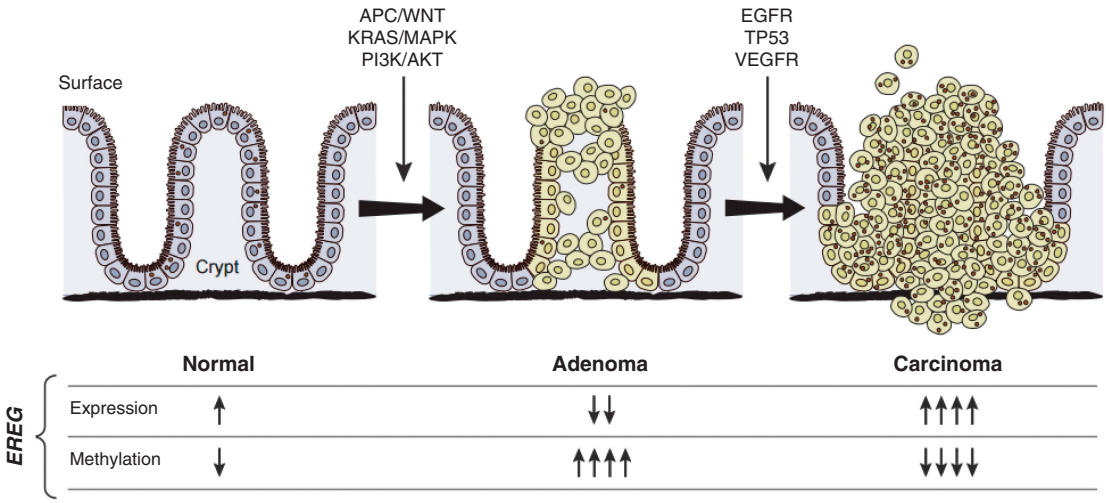

Figure 7. A model for EGFR activation during CRC progression through demethylation and subsequent upregulation of *EREG* at the adenoma–carcinoma transition. Colonic crypt and surface epithelial cells exhibit intermediate levels of methylation and low levels of *EREG* expression. The *EREG* promoter is heavily methylated in adenomas. This results in downregulation of *EGFR* and low levels of EGFR pathway activation. The EGFR pathway can be activated at the adenoma–carcinoma transition through demethylation of the *EREG* promoter and subsequent upregulation of *EREG* in a subset of CRCs. Black dots within cells represent an EREG transcript.

methylation and high levels of gene expression. It is likely that other mechanisms of *EREG* regulation exist beyond methylation, as we observed cases where *EREG* was upregulated in the absence of demethylation in some of our matched adenoma/carcinoma samples. Although AREG levels were higher in carcinomas compared with adenomas, the expression was not significantly associated with changes in the methylation status of the gene based on our data. The fact that *EREG* expression was low in normal colonic epithelium and adenomas, and high in carcinomas, while *AREG* was abundantly expressed in both malignant and non-malignant tissues, raises the possibility for different roles for *EREG* and *AREG* in the normal and malignant settings. This is consistent with reports on the unique role of EREG in the development of colitis-associated neoplasms and potentially other types of cancers.[50,53]

Activation of EGFR signaling in a subset of CRCs through *EREG* demethylation raises the question of whether epigenetic control is a mechanism for regulating other key pathways that drive CRC development. Interestingly, when we examined the clinically important VEGF pathway, we observed a striking increase in the expression of gene encoding its ligand, VEGFA, as well as a significant increase in the levels of VEGFR2 phosphorylation during the adenoma–carcinoma transition; however, this was not associated with significant changes in the methylation status of *VEGFA* (Supplementary Figure S10, data not shown). Hence, it is possible that epigenetic control of *EREG* might represent a targeted mechanism for EGFR regulation in CRC, rather than the by-product of global CIMP. In support of this notion, although we found a significant association between *EREG* methylation status and CIMP+ status in CRC samples from the TCGA collection, we found a subset of CIMP− tumors to exhibit high levels of methylation and low levels of *EREG* expression.

We confirmed the inverse correlation between *EREG* expression and methylation in CRC cell lines. Furthermore, we experimentally demonstrated that treatment of cell lines with high basal levels of methylation and low levels of *EREG* expression with 5-aza-dC resulted in reduced levels of methylation and enhanced expression of the gene. Although our demethylation studies were not gene-specific, our findings are consistent with the notion that reduced levels of *EREG* promoter methylation can lead to the transcriptional upregulation of the EGFR ligand at the adenoma to carcinoma transition. The phosphorylation levels of EGFR were also inversely correlated with *EREG* methylation in CRC cell lines. Demethylation experiments using 5-aza-dC in cell lines with high basal levels of *EREG* methylation led to a significant induction of *EREG* and a concomitant increase in EREG phosphorylation to levels typically observed in high *EREG*-expressing lines. Importantly, this also resulted in sensitization of CRC cells to the EGFR inhibitors, gefitinib and erlotinib, thus providing preliminary functional evidence of the role of epigenetic regulation of *EREG* as mechanism of EGFR pathway activation in CRC. In tissues, we also showed that upregulation of *EREG* was associated with a concomitant increase in the phosphorylation levels of the receptor during progression from adenoma to carcinoma. Studies have suggested that the ERK/MAPK pathway is activated and might help drive tumorigenesis downstream of EREG-driven EGFR phosphorylation in hepatocytes and in glioblastomas.[37,38] Although we observed higher levels of ERK/MAPK signaling in CRC compared with adenomas in our sample set, future studies will be required to assess the signaling downstream of EGFR-Y1068 and -Y1173 in CRC in more detail.

To begin to assess the clinical consequences of EREG demethylation, we examined the relationship between *EREG* promoter methylation and expression levels in tumors from patients treated with cetuximab as part of a phase II clinical trial. The fact that we observed a strong inverse correlation between *EREG* methylation and expression levels in these tissues is consistent with the notion of epigenetic regulation of *EREG* expression in CRC. When we assessed *EREG* promoter methylation status as it related to tumor measurements following cetuximab treatment, we found that tumors that exhibited the best CT responses after therapy were those with the lowest levels of *EREG* methylation. Tumors with the greatest size reductions following treatment expressed high levels of EREG. Conversely, tumors that grew following cetuximab treatment exhibited high levels of methylation and low levels of *EREG* expression. These data provide clinical support for a mechanism of EGFR pathway activation during CRC development through *EREG* promoter demethylation. Studies have previously shown that a positive CIMP status can lead to a diminished response to 5-fluorouracil-containing therapies in metastatic CRC.[41,42] Thus, it is possible that CIMP status did have an impact on response to treatment in our single arm cohort. However, we do not believe that this fully explains the association we observed between *EREG* methylation and clinical

response in our phase II study, in part, because *EREG* methylation did not always coincide with CIMP+ status, as we observed high levels of *EREG* methylation (and low levels of expression) in a subset of CIMP − samples from the TCGA collection. The EGFR antagonist, cetuximab, likely also had a role in contributing to the CT responses in our cohort. It is noteworthy that some patients from our phase II cohort whose tumors had high levels of *EREG* methylation (some of which would be expected to be CIMP+) did exhibit clinically meaningful CT responses. This may suggest that several predictive factors are at play, including EREG, CIMP status and potentially others that we are not aware of at this point in time.

Our finding that *EREG* expression can be epigenetically regulated in a subset of CRC raised the possibility that this mechanism might also be relevant in other cancer indications. We investigated the relationship between *EREG* promoter methylation and expression in available TCGA data sets. We found an inverse correlation between *EREG* methylation and expression levels in several types of malignancies. Most of these cancers were ones in which EGFR-targeting therapies have demonstrated clinical activity, such as neoplasms of the head and neck, lung and bladder.[40,45–48,54] Interestingly, not all of these cancer types exhibited this relationship between *EREG* methylation and expression, suggesting that EGFR pathway activation might proceed through other mechanisms in those tumor types. For example, increased methylation and inactivation of *EREG* is known to occur during the development of gastric cancer.[55]

In this study, we provide data that support a model for EGFR pathway activation during the evolution of CRC that may be driven through demethylation and subsequent upregulation of *EREG* at the adenoma–carcinoma transition in a subset of tumors (Figure 7). Future studies will be required to elucidate this potential mechanism of EGFR pathway activation as a driver of tumor progression in CRC and other cancer types.

MATERIALS AND METHODS

Patient and tissue samples

The use of the discovery set of samples from the Department of Pathology at the University of Virginia was approved by their Institutional Review Board. Laser microdissection was performed using a Leica AS LMD system (Leica Microsystems Inc., Bannockburn, IL, USA). Clinical details of the discovery set of 58 laser-captured frozen tissue samples and 48 macrodissected frozen CRC samples can be found in Supplementary Tables S1 and S2, respectively. To our knowledge, samples were obtained before therapy; however, 2/17 carcinomas and 1/17 adenomas were of rectal origin and patients may have received neoadjuvant therapy before tissue resection. The validation set of 16 matched adenomas and carcinomas were obtained from the MT Group (Van Nuys, CA, USA) and had appropriate Institutional Review Board approval. Clinical characteristics of the validation set are shown in Supplementary Table S5. Information on samples from the cetuximab arm of the phase II MEHD7945 trial (NCT01652482) can be found in Supplementary Table S8. All tissues were subjected to review by a pathologist to confirm the diagnosis and to define each cell type-enriched areas for macrodissection. Total RNA was purified using High Pure FFPE (Formalin-Fixed, Paraffin-Embedded) Micro Kit (Roche Diagnostics, Indianapolis, IN, USA). Total DNA was prepared by QIAamp DNA FFPE Tissue Kit (Qiagen, Valencia, CA, USA) and DNeasy Blood and Tissue Kit (Qiagen), respectively.

Next-generation sequencing

Mutations were detected using a previously developed MMP-Seq targeted cancer panel.[30] DNA sample quality was quantified as the number of functional copies using a TRAK2 qPCR 'ruler assay'.[30] Approximately 5000 functional copies of DNA from each sample were used as the input for target enrichment and library construction using Fluidigm Access Array (Fluidigm Corporation, South San Francisco, CA, USA), followed by deep sequencing on an Illumina MiSeq sequencer (Illumina, San Diego, CA, USA). Sequence alignment, primary variant calling and filtering were performed as described previously.[30]

Gene expression analysis

Analyses were performed using the R programming language (version 3.2). Gene expression profiles were collected on Affymetrix HG-U133A GeneChips (Affymetrix, Santa Clara, CA, USA) according to the manufacturer's instructions. Gene expression data was deposited into the GEO database under the accession number GSE77953. Gene expression values were obtained by quantile normalization and calculation of the robust multichip average expression measure using the *affy* Bioconductor package (version 1.46.1).[56] To account for potential confounding variables, for example, batch effects, surrogate variable analysis was performed with the *sva* Bioconductor package (version 3.14.0), which identified one variable that was included when fitting linear models (see below). For hierarchical clustering (Figure 1b), we fit the normalized gene expression values versus the surrogate variables and used the observed residuals to identify the top 500 genes with the largest interquartile range. Residual gene expression scores were transformed to *z*-scores and clustered based on 1–Pearson correlation as distance metric for Ward's clustering method. Principal component analysis was performed using residual gene expression scores from all assayed probes and the first two components were plotted (Supplementary Figure S1, PC1 and PC2).

We identified genes displaying differential expression between successive stages of CRC progression by applying a moderated *t*-test using the *limma* Bioconductor package (version 3.24.12). To account for potential confounding variables, the surrogate variable identified above was included together with the CRC progression stage in the linear model. Successive stages were contrasted and genes with false discovery rate < 0.05 and a minimum absolute log 2 fold change of $\geqslant 0.58$ (1.5-fold up- or downregulated) were considered significantly differentially expressed. Data were analyzed through the use of QIAGEN's Ingenuity Pathway Analysis (QIAGEN, Redwood City, CA, USA; www.qiagen.com/ingenuity). Genes that were differentially expressed between normal colonic surface epithelium compared with crypt cells or between carcinomas compared with adenomas at a log 2 fold change of $\geqslant 0.58$ were included in the pathway analysis (Supplementary Table S3). Pathway significance was measured as –log (*P*-value) and ranged from 0 to 38.5, and a list of pathways and associated significance is provided in Supplementary Table S4.

Taqman real-time PCR assay

The high-capacity cDNA Reverse Transcription Kit (cat. no. 4368814; Applied Biosystems, South San Francisco, CA, USA) was used to prepare cDNA from 200ng of total RNA. Relative cDNA quantification for *EREG* expression using *GAPDH* as an internal reference gene was carried out using Applied Biosystems ViiA7 Real-Time PCR System (Thermo Fisher Scientific, South San Francisco, CA, USA) following the standard protocol of the manufacturer. The primers and probe sets for each gene are from Life Technology's single tube gene expression assays. The catalog numbers for *EREG* and *GAPDH* are Hs00914313_m1 and Hs99999905_m1 (Life Technologies, South San Francisco, CA, USA).

Illumina infinium analysis

Genome-wide DNA methylation analysis was performed on DNAs from 48 fresh-frozen samples. Microarray data were collected at Expression Analysis Inc. (Durham, NC, USA) using the IlluminaHumanMethylation450 Beadchip (Illumina) and preprocessed using the Bioconductor *methylumi* software package (version 2.14.0; PMID 18467348) as described previously.[40] Genome-wide methylation data were deposited into the GEO database under the accession number GSE77954. Methylation values were reported as M-values (log 2 ratios of methylated to unmethylated probes). To determine probe-wise methylation scores, we associated each probe with the nearest annotated transcriptional start site, focusing on probes within the putative promoter regions of annotated genes (Supplementary Table S6).

To subgroup TCGA tumors into *MLH1* methylation high/low categories, we obtained raw Illumina 450k Beadchip microarray data for 292 tumor samples from the TCGA and preprocessed it as described above. As described previously, we examined probe cg00893636 located in the CpG island of the bidirectional MLH1/EPM2AIP1 promoter. The normalized methylation scores (M-value) were transformed to *z*-scores by centering on the mean and scaling to unit variance. As expected, visual inspection revealed a strongly bimodal distribution and samples with positive scores were assigned the MLH1/CIMP+ subgroup label.

Quantitative methylation-specific PCR

To quantify the percentage of methylated *EREG* at probe cg.19308222 in the validation sample set, CRC cell lines and in trial patient samples, quantitative methylation-specific PCR assays targeting either a fully methylated *EREG* (*EREG* qMSP), or a fully unmethylated *EREG* (*EREG* qUSP) at this specific position were designed. The sequences of real-time PCR primers for *EREG* quantitative methylation-specific PCR are as follows: *EREG* qMSP, forward, 5′-AGGGGTTTTAGAAGGAAGGC-3′, reverse, 5′-TATCA ATAATTCAAACGCCCTC-3′ and probe, 5′-TTACATTAAAACGCCACCGCCCAA-3′; *EREG* qUSP, forward, 5′-TAGGGGTTTTAGAAGGAAGGT-3′, reverse, 5′-TATC AATAATTCAAACACCCTCCTT-3′ and probe, 5′-CTTACATTAAAACACCACC ACCCAA-3′. DNA was isolated from FFPE tissues, and a previously developed preamplification procedure was applied to amplify the sodium bisulfite-converted DNA.[40] The preamplified material was then amplified in a second PCR reaction using Applied Biosystems ViiA7 Real-Time PCR System as described previously.[40] The quantification of methylated EREG was calculated as a ratio of methylated template to total input template and was based on standard curves generated using completely methylated and completely unmethylated DNA from Epi Tect (Qiagen).

5-Aza-dC treatment

All cell lines were obtained from the Genentech cell lines repository and were authenticated by STR profiling and confirmed to be free of mycoplasma contamination. Cells were grown in RPMI-1640 supplemented with 10% fetal bovine serum and 2 mM L-glutamine. Cells were seeded on day 0 at 4000–9000 cells per cm^2 and dosed with 500 nM 5-aza-2′ deoxycytidine (5-Aza-dC) (cat. no.: A3656; Sigma-Aldrich, St Louis, MO, USA) or dimethyl sulfoxide (DMSO) control on days 1 and 3. On day 5, cells were washed once with cold phosphate-buffered saline and harvested by scraping in RLT plus buffer (Qiagen). DNA and RNA were isolated using AllPrep DNA/RNA Mini Kit (Qiagen). *EREG* expression level was quantified using real-time PCR as described above. The *EREG* methylation at probe cg.19308222 was measured and calculated using the quantitative methylation-specific PCR assays described above. Total EGFR and phospho-EGFR levels were measured by western blot using the same antibodies as RPPA assay. Cells were starved with serum-free RPMI-1640 with 2 mM glutamine for 4 h before the cells were harvested and lysed for western blot. Western antibody for β-actin (cat. no.: sc-47778) was obtained from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Western antibodies for total EGFR (cat. no.: 4267), phospho-EGFR-1068 (cat. no.: 3777) and phospho-EGFR-Y1173 (cat. no.: 4407) were all purchased from Cell Signaling Technology (Danvers, MA, USA).

Gefitinib and erlotinib treatment

To determine the effect of 5-aza-dC on the sensitivity of CRC cell lines to gefitinib and erlotinib treatment, cells were pretreated with 500 nM 5-aza-dC for 2 days. Pretreated cells were plated in quadruplicate in 384-well plates in RPMI containing 10% fetal bovine serum, and incubated overnight. The cells were then treated with different concentrations of gefitinib or erlotinib alone or in combination with 250 nM 5-aza-dC. After 72 h, cell viability was measured using the CellTiter-Glo Luminescent Cell Viability assay (Promega, Sunnyvale, CA, USA). Cell viability for gefitinib or erlotinib treatment alone was normalized to DMSO-treated control cells, whereas cell viability for gefitinib or erlotinib treatment combined with 5-aza-dC was normalized to DMSO plus 250 nM 5-aza-dC-treated control cells. To evaluate the effect of 5-aza-dC treatment on the sensitivity of CRC cell lines to gefitinib and erlotinib treatments, an unpaired *t*-test was used and a two-sided *P*-value was derived. Erlotinib (cat. no.: G00022086.23-1) was obtained from Mosaic Life Care (St Joseph, MO, USA), and gefitinib (cat. no.: 2030-5) was purchased from BioVision (Milpitas, CA, USA).

RPPA analysis

RPPA was performed by Theranostics Health (Rockville, MD, USA) according to the manufacturer's instruction. Protein lysates were extracted from pulverized fresh-frozen samples. Four replicates for three dilutions from each lysate were printed on the slides; the dilutions aided in total protein normalization and confirmation of the linear range of detection. Printed slides were subjected to immunostaining with antibodies purchased from Cell Signaling Technology, and validated for RPPA by Theranostics Health. The raw expression data in each spot was normalized to the ss-DNA content, in which the data have been corrected for potential spatial and total protein bias in the raw data.

RNA-ISH

RNA-ISH for Hs-EREG was performed using RNAscope 2.0 FFPE Reagent Kit by Advanced Cell Diagnostics (Hayward, CA, USA) according to the manufacturer's instructions. Briefly, 5 μm FFPE tissue sections were pretreated with heat and protease before hybridization with the target oligo probe (NM_001432.2, nt. 253–2317). Preamplifier, amplifier and horse radish peroxidase-labeled oligos were then hybridized sequentially, followed by chromogenic precipitate development. Specific RNA staining signal was identified as brown, punctate dots. Each sample was quality controlled for RNA integrity with an RNAscope probe specific to PPIB RNA and for background with a probe specific to bacterial dapB RNA. To verify technical and scoring accuracy, reference slides consisting of FFPE HeLa cell pellets were tested for PPIB and dapB together with tissue FFPE slides. The area of interest was reviewed at x20 magnification. RNAscope signal was binned into five groups based on the number of dots per cell as follows: bin 0 = 0 dots per cell, bin 1 = 1–3 dots per cell, bin 2 = 4–9 dots per cell, bin 3 = 10–15 dots per cell and bin 4 = > 15 dots per cell with > 10% of dots in clusters. Each sample was evaluated for the percentage of cells in each bin. The H-score was calculated as follows, H-score = (0 x (% of cells in bin 0)) + (1 x (% of cells in bin 1)) + (2 x (% of cells in bin 2)) + (3 x (% of cells in bin 3)) + (4 x (% of cells in bin 4)), with an H-score scale of 0–400. Bright field images were acquired using a Zeiss Axio Imager M1 microscope (Zeiss, Oberkochen, Germany) using a ×40 objective.

Statistical analysis

To evaluate the correlation between normalized mRNA expression and mean promoter methylation (M-value) during CRC progression, Spearman's rank-correlation coefficients were calculated for all genes that were (1) assayed on the gene expression arrays and (2) whose promoter(s) contained one or more probes assayed on the methylation arrays. In the case of genes for which methylation probes mapped to multiple alternative promoters, the promoter displaying the highest interquartile range across the full data set was selected. To evaluate the differential expression of annotated EGFR and VEGFR signaling pathway components during CRC progression, an unpaired *t*-test was used and two-sided *P*-value was derived. To compare *EREG* expression and promoter methylation differences between adenoma and carcinoma in the validation sample set, a paired *t*-test was applied and the two-sided *P*-value was derived. To assess the statistical significance of the correlation among % best CT response, *EREG* methylation and *EREG* expression, we performed a simple linear regression analysis and reported the associated *P*-values.

CONFLICT OF INTEREST

XQ, TS, LF, KW, KO, SL, YW, RB, EP, AP, LA, MRL, GMH and OK are current or former employees and stock holders of Genentech/Roche. The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank Dr Ron Firestein for pathology review, Shoji Ikeda for helping with the acquisition of clinical tissues, Cheryl Wong, An Do, Anna Cheung, My Vo and Dr Teiko Sumiyoshi for assistance with the tissue processing and RNA and DNA preparation, and Doris Kim and Ling-Yuh Huw for helpful discussions. We thank Allison Bruce for assistance with the CRC progression model illustration, and A. Daisy Goodrich for editorial review of the manuscript. The authors also acknowledge the *EREG* expression and methylation data across different indications generated by the TCGA Research Network: http://cancergenome.nih.gov. The authors also thank the MEHD7945A trial team, and all the patients who participated in the trial.

REFERENCES

1 Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759–767.
2 de Santa Barbara P, van den Brink GR, Roberts DJ. Development and differentiation of the intestinal epithelium. *Cell Mol Life Sci* 2003; **60**: 1322–1332.
3 de Lau W, Barker N, Clevers H. WNT signaling in the normal intestine and colorectal cancer. *Front Biosci* 2007; **12**: 471–491.

4 Gryfe R, Swallow C, Bapat B, Redston M, Gallinger S, Couture J. Molecular biology of colorectal cancer. *Curr Probl Cancer* 1997; **21**: 233–300.
5 Nathke IS. The adenomatous polyposis coli protein: the Achilles heel of the gut epithelium. *Annu Rev Cell Dev Biol* 2004; **20**: 337–366.
6 Oving IM, Clevers HC. Molecular causes of colon cancer. *Eur J Clin Invest* 2002; **32**: 448–457.
7 Pasparakis M. Role of NF-kappaB in epithelial biology. *Immunol Rev* 2012; **246**: 346–358.
8 Skeen VR, Paterson I, Paraskeva C, Williams AC. TGF-beta1 signalling, connecting aberrant inflammation and colorectal tumorigenesis. *Curr Pharm Des* 2012; **18**: 3874–3888.
9 van Es JH, Clevers H. Notch and Wnt inhibitors as potential new drugs for intestinal neoplastic disease. *Trends Mol Med* 2005; **11**: 496–502.
10 Watt FM. Unexpected Hedgehog–Wnt interactions in epithelial differentiation. *Trends Mol Med* 2004; **10**: 577–580.
11 Grady WM, Markowitz SD. Genetic and epigenetic alterations in colon cancer. *Annu Rev Genom Hum Genet* 2002; **3**: 101–128.
12 Luo Y, Wong CJ, Kaz AM, Dzieciatkowski S, Carter KT, Morris SM *et al.* Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology* 2014; **147**: e418.
13 Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D. Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer* 2009; **9**: 489–499.
14 Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487**: 330–337.
15 Yao YL, Shao J, Zhang C, Wu JH, Zhang QH, Wang JJ *et al.* Proliferation of colorectal cancer is promoted by two signaling transduction expression patterns: ErbB2/ErbB3/AKT and MET/ErbB3/MAPK. *PLoS One* 2013; **8**: e78086.
16 Douillard JY, Siena S, Cassidy J, Tabernero J, Burkes R, Barugel M *et al.* Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J Clin Oncol* 2010; **28**: 4697–4705.
17 Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W *et al.* Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004; **350**: 2335–2342.
18 Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A *et al.* Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009; **360**: 1408–1417.
19 Lievre A, Bachet JB, Boige V, Cayre A, Le Corre D, Buc E *et al.* KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *J Clin Oncol* 2008; **26**: 374–379.
20 Peeters M, Douillard JY, Van Cutsem E, Siena S, Zhang K, Williams R *et al.* Mutant KRAS codon 12 and 13 alleles in patients with metastatic colorectal cancer: assessment as prognostic and predictive biomarkers of response to panitumumab. *J Clin Oncol* 2013; **31**: 759–765.
21 Pentheroudakis G, Kotoula V, De Roock W, Kouvatseas G, Papakostas P, Makatsoris T *et al.* Biomarkers of benefit from cetuximab-based therapy in metastatic colorectal cancer: interaction of EGFR ligand expression with RAS/RAF, PIK3CA genotypes. *BMC Cancer* 2013; **13**: 49.
22 Price TJ, Hardingham JE, Lee CK, Weickhardt A, Townsend AR, Wrin JW *et al.* Impact of KRAS and BRAF gene mutation status on outcomes from the phase III AGITG MAX Trial of capecitabine alone or in combination with bevacizumab and mitomycin in advanced colorectal cancer. *J Clin Oncol* 2011; **29**: 2675–2682.
23 Di Nicolantonio F, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P *et al.* Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* 2008; **26**: 5705–5712.
24 Khambata-Ford S, Garrett CR, Meropol NJ, Basik M, Harbison CT, Wu S *et al.* Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol* 2007; **25**: 3230–3237.
25 Jonker DJ, Karapetis CS, Harbison C, O'Callaghan CJ, Tu D, Simes RJ *et al.* Epiregulin gene expression as a biomarker of benefit from cetuximab in the treatment of advanced colorectal cancer. *Br J Cancer* 2014; **110**: 648–655.
26 Jones S, Chen WD, Parmigiani G, Diehl F, Beerenwinkel N, Antal T *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci USA* 2008; **105**: 4283–4288.
27 Makky K, Tekiela J, Mayer AN. Target of rapamycin (TOR) signaling controls epithelial morphogenesis in the vertebrate intestine. *Dev Biol* 2007; **303**: 501–513.
28 Sancho E, Batlle E, Clevers H. Signaling pathways in intestinal development and cancer. *Annu Rev Cell Dev Biol* 2004; **20**: 695–723.
29 Fearon ER. Molecular genetics of colorectal cancer. *Annu Rev Pathol* 2011; **6**: 479–507.
30 Bourgon R, Lu S, Yan Y, Lackner MR, Wang W, Weigman V *et al.* High-throughput detection of clinically relevant mutations in archived tumor samples by multiplexed PCR and next-generation sequencing. *Clin Cancer Res* 2014; **20**: 2080–2091.
31 Zhang H, Berezov A, Wang Q, Zhang G, Drebin J, Murali R *et al.* ErbB receptors: from oncogenes to targeted cancer therapies. *J Clin Invest* 2007; **117**: 2051–2058.
32 Yonesaka K, Zejnullahu K, Lindeman N, Homes AJ, Jackman DM, Zhao F *et al.* Autocrine production of amphiregulin predicts sensitivity to both gefitinib and cetuximab in EGFR wild-type cancers. *Clin Cancer Res* 2008; **14**: 6963–6973.
33 Oshima G, Wennerberg J, Yamatodani T, Kjellen E, Mineta H, Johnsson A *et al.* Autocrine epidermal growth factor receptor ligand production and cetuximab response in head and neck squamous cell carcinoma cell lines. *J Cancer Res Clin Oncol* 2012; **138**: 491–499.
34 Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr LA, Kinzler KW. Cancer genome landscapes. *Science (New York, NY)* 2013; **339**: 1546–1558.
35 Akiyama Y, Watkins N, Suzuki H, Jair KW, van Engeland M, Esteller M *et al.* GATA-4 and GATA-5 transcription factor genes and potential downstream antitumor target genes are epigenetically silenced in colorectal and gastric cancer. *Mol Cell Biol* 2003; **23**: 8429–8439.
36 Orre LM, Panizza E, Kaminskyy VO, Vernet E, Graslund T, Zhivotovsky B *et al.* S100A4 interacts with p53 in the nucleus and promotes p53 degradation. *Oncogene* 2013; **32**: 5531–5540.
37 Kohsaka S, Hinohara K, Wang L, Nishimura T, Urushido M, Yachi K *et al.* Epiregulin enhances tumorigenicity by activating the ERK/MAPK pathway in glioblastoma. *Neuro-oncology* 2014; **16**: 960–970.
38 Auf G, Jabouille A, Delugin M, Guerit S, Pineau R, North S *et al.* High epiregulin expression in human U87 glioma cells relies on IRE1alpha and promotes autocrine growth through EGF receptor. *BMC Cancer* 2013; **13**: 597.
39 Komurasaki T, Toyoda H, Uchida D, Nemoto N. Mechanism of growth promoting activity of epiregulin in primary cultures of rat hepatocytes. *Growth Factors (Chur, Switzerland)* 2002; **20**: 61–69.
40 Walter K, Holcomb T, Januario T, Du P, Evangelista M, Kartha N *et al.* DNA methylation profiling defines clinically relevant biological subsets of non-small cell lung cancer. *Clin Cancer Res* 2012; **18**: 2360–2373.
41 Jover R, Nguyen TP, Perez-Carbonell L, Zapater P, Paya A, Alenda C *et al.* 5-Fluorouracil adjuvant chemotherapy does not increase survival in patients with CpG island methylator phenotype colorectal cancer. *Gastroenterology* 2011; **140**: 1174–1181.
42 Shen L, Catalano PJ, Benson AB III, O'Dwyer P, Hamilton SR, Issa JP. Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy. *Clin Cancer Res* 2007; **13**: 6093–6098.
43 Seniski GG, Camargo AA, Ierardi DF, Ramos EA, Grochoski M, Ribeiro ES *et al.* ADAM33 gene silencing by promoter hypermethylation as a molecular marker in breast invasive lobular carcinoma. *BMC Cancer* 2009; **9**: 80.
44 Weng WH, Leung WH, Pang YJ, Hsu HH. Lauric acid can improve the sensitization of Cetuximab in KRAS/BRAF mutated colorectal cancer cells by retrievable microRNA-378 expression. *Oncol Rep* 2016; **35**: 107–116.
45 Sun JZ, Lu Y, Xu Y, Liu F, Li FQ, Wang QL *et al.* Epidermal growth factor receptor expression in acute myelogenous leukaemia is associated with clinical prognosis. *Hematol Oncol* 2012; **30**: 89–97.
46 Riese DJ 2nd, Cullum RL. Epiregulin: roles in normal physiology and cancer. *Semin Cell Dev Biol* 2014; **28**: 49–56.
47 Deangelo DJ, Neuberg D, Amrein PC, Berchuck J, Wadleigh M, Sirulnik LA *et al.* A phase II study of the EGFR inhibitor gefitinib in patients with acute myeloid leukemia. *Leuk Res* 2014; **38**: 430–434.
48 Lainey E, Wolfromm A, Marie N, Enot D, Scoazec M, Bouteloup C *et al.* Azacytidine and erlotinib exert synergistic effects against acute myeloid leukemia. *Oncogene* 2013; **32**: 4331–4342.
49 Hansen AR, Siu LL. Epidermal growth factor receptor targeting in head and neck cancer: have we been just skimming the surface? *J Clin Oncol* 2013; **31**: 1381–1383.
50 Busser B, Sancey L, Brambilla E, Coll JL, Hurbin A. The multiple roles of amphiregulin in human cancer. *Biochim Biophys Acta* 2011; **1816**: 119–131.
51 Revillion F, Lhotellier V, Hornez L, Bonneterre J, Peyrat JP. ErbB/HER ligands in human breast cancer, and relationships with their receptors, the bio-pathological features and prognosis. *Ann Oncol* 2008; **19**: 73–80.
52 Troiani T, Martinelli E, Napolitano S, Vitagliano D, Ciuffreda LP, Costantino S *et al.* Increased TGF-alpha as a mechanism of acquired resistance to the anti-EGFR inhibitor cetuximab through EGFR-MET interaction and activation of MET signaling in colon cancer cells. *Clin Cancer Res* 2013; **19**: 6751–6765.

53 Neufert C, Becker C, Türeci Ö, Waldner MJ, Backert I, Floh K *et al*. Tumor fibroblast-derived epiregulin promotes growth of colitis-associated neoplasms through ERK. *J Clin Invest* 2013; **123**: 1428–1443.

54 Rebouissou S, Bernard-Pierrot I, de Reynies A, Lepage ML, Krucker C, Chapeaublanc E *et al*. EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype. *Sci Transl Med* 2014; **6**: 244–291.

55 Yun J, Song SH, Park J, Kim HP, Yoon YK, Lee KH *et al*. Gene silencing of EREG mediated by DNA methylation and histone modification in human gastric cancers. *Lab Invest* 2012; **92**: 1033–1044.

56 Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003; **31**: e15.

Supplementary Information accompanies this paper on the Oncogene website (http://www.nature.com/onc)