HHS Public Access

Author manuscript

Nat Med. Author manuscript; available in PMC 2018 May 21.

Published in final edited form as:

Nat Med. 2017 November ; 23(11): 1331–1341. doi:10.1038/nm.4424.

lncRNA MIR100HG-derived miR-100 and miR-125b mediate cetuximab resistance via Wnt/β-catenin signaling

Yuanyuan Lu[1,2,*], **Xiaodi Zhao**[1,2,*], **Qi Liu**[3], **Cunxi Li**[4], **Ramona Graves-Deal**[1], **Zheng Cao**[1], **Bhuminder Singh**[1], **Jeffrey L. Franklin**[1], **Jing Wang**[3], **Huaying Hu**[4], **Tianying Wei**[4], **Mingli Yang**[5], **Timothy J. Yeatman**[5], **Ethan Lee**[6], **Kenyi Saito-Diaz**[6], **Scott Hinger**[7], **James G. Patton**[7], **Christine H. Chung**[8], **Stephan Emmrich**[9], **Jan-Henning Klusmann**[9], **Daiming Fan**[2,§], and **Robert J. Coffey**[1,10,§]

[1]Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA

[2]State Key Laboratory of Cancer Biology, National Clinical Research Center for Digestive Diseases and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Xi'an, Shaanxi 710032, China

[3]Department of Biomedical Informatics and Center for Quantitative Sciences, Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA

[4]Jiaen Genetics Laboratory, Beijing Jiaen Hospital, Beijing and Molecular Pathology, Cancer research center, Medical College of Xiamen University, Xiamen, Fujian 361102, China

[5]Gibbs Cancer Center & Research Institute, Spartanburg, South Carolina 29303, USA

[6]Department of Cell and Developmental Biology and Vanderbilt Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

[7]Department of Biological Sciences, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

[8]Moffitt Cancer Center & Research Institute, Tampa, Florida 33612, USA

[9]Pediatric Hematology and Oncology, Hannover Medical School, Hannover 30625, Germany

[10]Department of Veterans Affairs Medical Center, Nashville, Tennessee 37212, USA

[§]Address correspondence to: Robert J. Coffey, MD, Epithelial Biology Center, 10415 MRB IV, Vanderbilt University Medical Center, 2213 Garland Ave, Nashville, TN 37232, Fax: (615) 343-1591, robert.coffey@vanderbilt.edu; Daiming Fan, MD, PhD, Xijing Hospital of Digestive Diseases, Fourth Military Medical University, 127 Changle West Road, Xi'an, Shaanxi 710032, China, Fax: (86) 29-82539041, fandaim@fmmu.edu.cn.

[*]These authors contributed equally to this work.

Author Contributions: Y.L., X.Z., C.L., D.F. and R.J.C. designed the research. Y.L., X.Z., Q.L., C.L., R.G.-D., Z.C., B.S., J.W., H.H., T.W., M.Y. and S.H. performed experiments, analyzed data and prepared figures and tables. T.Y., E.L., K.S.-D., C.H.C., S.E., J.-H.K. and D.F. contributed to new reagents and/or analytical tools. Y.L., X.Z., Q.L., C.L., J.L.F., T. J.Y., E.L., J.G.P., C.H.C., D.F. and R.J.C. analyzed the data and provided critical input. Y.L., X.Z. and R.J.C. wrote the paper. R.J.C. and D.F. conceived the project, supervised and coordinated all aspects of the work.

Competing Financial Interests Statement: The authors declare no competing financial interests.

Abstract

De novo and acquired resistance, largely attributed to genetic alterations, are barriers to effective anti-EGFR therapy. We generated cetuximab-resistant cells following prolonged cetuximab exposure to cetuximab-sensitive colorectal cancer cells in three-dimensional culture. Through whole exome sequencing and transcriptional profiling, we found overexpression of lncRNA MIR100HG and two embedded miRNAs, miR-100 and miR-125b, in the absence of known genetic events linked to cetuximab resistance. MIR100HG and miR-100/125b overexpression was also observed in cetuximab-resistant colorectal cancer and head and neck squamous cell cancer cell lines and in tumors from colorectal cancer patients that progressed on cetuximab. miR-100/125b coordinately represses five Wnt/β-catenin negative regulators, resulting in increased Wnt signaling, and Wnt inhibition in cetuximab-resistant cells restored cetuximab responsiveness. We describe a double-negative feedback loop between MIR100HG and GATA6, whereby GATA6 represses MIR100HG, but this repression is relieved by miR-125b targeting of GATA6. These studies identify a clinically actionable, epigenetic cause of cetuximab resistance.

Colorectal cancer (CRC) remains a leading cause of cancer-related death worldwide[1]. Cetuximab and panitumumab are EGF receptor (EGFR) monoclonal antibodies (mAbs) that bind the extracellular domain of EGFR and enhance receptor internalization and degradation. These EGFR mAbs are common targeted agents for patients with wild-type *KRAS* metastatic CRC. As monotherapy, 12–17% patients have durable responses[2] and up to 72% response rates are reported when combined with chemotherapy[3]. However, drug resistance frequently arises. Intense efforts have led to identification of many *de novo* and acquired genetic mechanisms of resistance to EGFR mAb therapy, including *KRAS*, *NRAS*, *BRAF*, *PIK3CA*, and *EGFR* mutations[2,4,5]. However, little is known about non-genetic resistance mechanisms.

Non-coding RNAs (ncRNAs), in particular long non-coding RNAs (lncRNAs) and microRNAs (miRNAs), play crucial roles in epigenetic regulation[6,7]. Recently, a complex interplay between these two classes of regulatory ncRNAs has been discovered in which some lncRNAs are processed to produce miRNAs that repress target mRNAs[8,9]. For example, lncRNA H19-derived miR-675 suppresses translation of insulin growth factor receptor (Igf1r), inhibiting cell proliferation in response to cellular stress or oncogenic signals[10]. miR-17~92, generated from the lncRNA MIR17HG locus, attenuates TGF-β signaling to stimulate angiogenesis and tumor growth[11]. The lncRNA MIR100HG-derived miR-100/let-7a-2/miR-125b-1 and MIR99AHG-derived miR-99a/let-7c/miR-125b-2 clusters participate in the pathogenesis of acute megakaryoblastic leukemia[12,13]. However, whether these lncRNAs or derived miRNAs contribute to drug resistance is largely unknown.

Herein, we identify a role for lncRNA MIR100HG and two embedded miRNAs, miR-100 and miR-125b, in conferring cetuximab resistance. We show that MIR100HG and miR-100/125b are overexpressed in the setting of *de novo* and acquired cetuximab resistance in CRC and head and neck squamous cell cancer (HNSCC) cell lines. miR-100 and miR-125b coordinately downregulate five negative regulators (DKK1, DKK3, ZNRF3, RNF43, APC2) of canonical Wnt/β-catenin signaling (hereafter Wnt signaling), leading to increased Wnt signaling. Wnt inhibition restores responsiveness to cetuximab *in vitro* and *in*

vivo. We show these events occur in CRC patients whose tumors progressed on cetuximab. We also discovered that MIR100HG overexpression is reinforced by miR-125b suppression of GATA6, which in turn represses MIR100HG. These studies identify an epigenetic cause of cetuximab resistance with diagnostic and therapeutic implications.

Results

Establishment of cetuximab-resistant cells in three-dimensional (3D) culture

By placing single cells from a human *KRAS/NRAS/BRAF* wild-type, microsatellite unstable CRC cell line, HCA-7, into 3D culture in type-1 collagen, a line was derived from colonies with cystic morphology and designated cystic colonies (CC) [14,15]. Proliferation of CC was inhibited by cetuximab in 3D culture but not in 2D plastic culture[14]. Upon continuous exposure to cetuximab in 3D culture for approximately 4 months, a line was generated and designated CC-cetuximab resistant (CC-CR) (Fig. 1a). In 2D culture, CC and CC-CR were morphologically indistinguishable. In 3D, however, CC formed hollow cysts with a central lumen lined by a monolayer of polarized cells, whereas CC-CR formed solid disorganized colonies (Fig. 1b). As expected, cetuximab inhibited CC growth in 3D, while CC-CR remained refractory to cetuximab up to 200 μg/ml (Fig. 1c; Extended Data Fig. 1a and b). Decreased expression of the proliferative marker, Ki-67, and increased expression of the apoptotic marker, cleaved Caspase-3, were observed in CC 24 h after cetuximab treatment, but these indices were unaffected in CC-CR (Fig. 1d). In cetuximab-treated CC, we observed reduced levels of p-EGFR, p-ERK1/2, p-AKT and Cyclin D1, as well as increased cleaved Caspase-3 and the pro-apoptotic marker, BIM; these markers were largely unaffected in cetuximab-treated CC-CR (Fig. 1e). Next, CC and CC-CR were stably transduced with a green fluorescent protein (GFP)-expressing lentiviral vector and injected subcutaneously into athymic nude mice. CC tumors were well differentiated and regressed upon administration of cetuximab. In contrast, CC-CR tumors were poorly differentiated and continued to grow in the presence of cetuximab, although not to the extent of untreated tumors (Fig. 1f and g; Extended Data Fig. 1c-e).

Upregulation of lncRNA MIR100HG and embedded miR-100/125b in cetuximab-resistant cells

We first considered known mechanisms of cetuximab resistance in this 3D model. By whole exome sequencing and RNA Sequencing (RNA-Seq), no known genetic events linked to cetuximab resistance were found, including all reported gene mutations, copy number changes and gene fusion events (Extended Data Table 1). By RNA-Seq, we found 141 transcripts upregulated and 220 transcripts downregulated in CC-CR compared to CC (fold change>2 and false-discovery rate, FDR<0.01). Expression levels of ERBB1-4, the 7 EGFR ligands, and MET were comparable between CC and CC-CR (Extended Data Table 2). Immunofluorescence also showed equivalent cell-surface EGFR staining in CC and CC-CR (Extended Data Fig. 1f). Small RNA-Seq detected 7 miRNAs upregulated and 24 miRNAs downregulated in CC-CR compared to CC (fold change>2 and FDR<0.01). Of note, the most upregulated transcript in CC-CR was lncRNA MIR100HG, and the two most upregulated miRNAs were miR-125b and miR-100 (Fig. 2a).

MIR100HG is the host gene of the miR-100/let-7a-2/miR-125b-1 cluster on chromosome 11 (Fig. 2b). qRT-PCR analysis confirmed upregulation of endogenous MIR100HG expression in CC-CR in the presence or absence of cetuximab (Fig. 2c). pri-miR-100, pri-miR-125b-1, and their corresponding mature miRNA, miR-100 and miR-125b, were also enriched in CC-CR (Fig. 2c and Extended Data Fig. 2a). Although pri-let-7a-2 was upregulated in CC-CR, mature let-7a was unchanged compared to CC (Extended Data Fig. 2b). The transcriptional start site (TSS) of MIR100HG was confirmed by 5′ RACE-PCR (Extended Data Fig. 2c). Analysis of The Cancer Genome Atlas (TCGA) CRC data repository revealed that miR-100 and miR-125b expression is tightly correlated with MIR100HG expression (Fig. 2d). RNA fluorescence *in situ* hybridization (FISH) showed highly enriched MIR100HG and miR-100/125b expression in CC-CR tumor xenografts (Fig. 2e). In contrast, let-7a expression did not correlate with that of MIR100HG (Extended Data Fig. 2d).

To assess whether MIR100HG and miR-100/125b overexpression extended beyond this one cell line, we examined their expression in a panel of 30 CRC cell lines placed upon a continuum of cetuximab sensitivity and resistance based upon published reports[16,17] (Extended Data Table 3). Expression of MIR100HG and miR-100/125b were enriched in more cetuximab-resistant lines compared to the more sensitive lines (Fig. 2f). Their expression inversely correlated with cetuximab resistance, regardless of *KRAS*/*BRAF* mutational status (Extended Data Fig. 2e and f). For example, two of the cetuximab-sensitive lines (GEO and SW403) expressed low levels of MIR100HG and miR-100/125b despite harboring mutant *KRAS.* In addition, we also observed upregulation of MIR100HG and miR-100/125b in the setting of cetuximab resistance in HNSCC cell lines (Extended Data Fig. 3a). Thus, MIR100HG and miR-100/125b are upregulated in the setting of cetuximab resistance in CRC and HNSCC cell lines and this phenomenon occurs in both acquired and *de novo* resistance. These findings led us to further explore the function of MIR100HG and miR-100/125b in cetuximab resistance.

miR-100 and miR-125b cooperativity drives cetuximab resistance

Since a major role of certain lncRNAs is production of embedded miRNAs[10,18], we asked whether cetuximab resistance is mediated by miR-100 and miR-125b overexpression. To this end, we delivered lentiviral-based overexpression or sponge constructs into CC and CC-CR, respectively, to generate stable cell lines expressing each miRNA, the miR-100/125b bicistron, or their corresponding sponges (Extended Data Fig. 3b and c). Although the miR-100 sponge had no significant effect on colony number in CC-CR in 3D culture, both the miR-125b and bicistron sponges significantly reduced colony number (Fig. 3a). In the presence of cetuximab, the miR-100 sponge modestly reduced colony number, whereas the reduction in colony number was more pronounced with the miR-125b sponge and the bicistron sponge (Fig. 3a). In contrast, opposite effects were observed in CC upon overexpressing miR-100 and miR-125b individually and together. The miR-100/125b bicistron, but not individual miRNAs, increased colony number in CC (Fig. 3b). Upon cetuximab treatment, the miR-100/125b bicistron conferred the strongest pro-survival effect; when introduced individually, miR-125b had a greater effect than miR-100 (Fig. 3b). Similar opposing effects were observed in morphological changes, as well as Ki-67 and cleaved Caspase-3 staining upon expressing the different sponges in CC-CR and the different

miRNAs in CC (Fig. 3c and d; Extended Data Fig. 3d and e). Additionally, overexpression of the miR-100/125b bicistron in Caco-2 cells (low endogenous miR-100/125b expression) rendered cells less responsive to cetuximab, whereas inhibition of the miR-100/125b bicistron in DLD-1 cells and HNSCC SCC25-derived CTX-R7 cells (both with high endogenous miR-100/125b expression) restored cetuximab responsiveness (Extended Data Fig. 4). Similar results were observed when CC-CR and CC cells with the differing manipulations were established as subcutaneous xenografts in nude mice and treated with cetuximab (Fig. 3e and f; Extended Data Fig. 5). Together, these results are consistent with a model in which miR-100 and miR-125b cooperate to confer cetuximab resistance.

miR-100 and miR-125b repress multiple Wnt negative regulators and increase Wnt signaling

To understand how miR-100 and miR-125b influence cetuximab responsiveness, we considered the most downregulated genes in CC-CR to be potential targets of these miRNAs. Two negative regulators of Wnt signaling, DKK1 and DKK3, were decreased over 30-fold in CC-CR compared to CC (Fig. 2a). Meanwhile, Wnt activity was enhanced in CC-CR by analysis of 64 consensus β-catenin target genes[19] (Extended Data Fig. 6a). Analysis of a large human CRC dataset (n=458) also revealed that MIR100HG expression levels positively correlated with the Wnt score[19], whereas no correlation was observed between MIR100HG and the Ras-Az score[20], which measures MEK activation as a downstream index of RAS signaling (Extended Data Fig. 6a and data not shown). Functional enrichment analysis further identified Wnt pathway enrichment in miR-100 and miR-125b putative targets (Extended Data Table 4). We thus considered whether miR-100 and miR-125b might target components of Wnt signaling. Through computational target prediction, we found that 3′ UTRs of DKK1 and DKK3 contain binding sites for miR-100 and miR-125b, respectively (Fig. 4a). Since clustered miRNAs are co-expressed and often coordinately regulate molecular pathways by targeting different components of the same pathway[21], we searched for other negative regulators of Wnt signaling that contain putative binding sites for miR-100 or miR-125b, and identified zinc and ring finger 3 (ZNRF3), ring finger protein 43 (RNF43), and APC2 as potential targets of miR-100 or miR-125b alone or in combination (Fig. 4a and Extended Data Table 5). Decreased protein levels of these five Wnt negative regulators in CC-CR compared to CC were confirmed by both immunoblots in cell lines and immunostaining in xenografts (Extended Data Fig. 6b and c). Using 3′ UTR luciferase reporter assays, we confirmed these five candidates are direct targets of miR-100 and/or miR-125b in both CC and Caco-2 cells; repression of these genes was rescued by mutations in the corresponding binding sites (Fig. 4b and c; Extended Data Fig. 6d). Immunoblots confirmed the regulation of the predicted targets by miR-100 and miR-125b alone or in combination in CC and CC-CR (Fig. 4d). Consistently, this regulation was also observed in Caco-2 cells (low endogenous miR-100/125b expression) and HuTu80 cells (high endogenous miR-100/125b expression) (Extended Data Fig. 6e).

We next examined whether miR-100/125b-induced downregulation of these Wnt negative regulators resulted in increased Wnt signaling. Although total β-catenin levels were not significantly altered, CC-CR exhibited increased active tyrosine phosphorylated p-Y489 β-catenin and increased nuclear β-catenin compared to CC in 3D (Fig. 4e and Extended Data

Fig. 6f). Consistently, β-catenin was largely confined to the plasma membrane in CC xenografts, whereas it was largely nuclear in CC-CR xenografts (Fig. 4f). Moreover, mRNA expression of a panel of Wnt target genes was significantly enriched in CC-CR versus CC (Fig. 4g). Cetuximab blocked Wnt3a-induced Wnt activation in CC, but had no obvious effect on Wnt3a-induced Wnt signaling in CC-CR (Extended Data Fig. 6g). Cetuximab also led to a marked and persistent decrease in Wnt target genes in CC over 48 h, whereas expression of those genes in CC-CR was only modestly decreased at early time points after treatment before rebounding at later time points (Extended Data Fig. 6h). Furthermore, nuclear β-catenin levels increased in CC and Caco-2 cells stably overexpressing either miR-100 or miR-125b, and the increase was greater in cells expressing the miR-100/125b bicistron (Extended Data Fig. 7a). In contrast, nuclear β-catenin levels were reduced upon overexpressing the miR-100/125b bicistron sponge in CC-CR, DLD-1, and CTX-R7 cells (Extended Data Fig. 7a). Corresponding changes of Wnt target genes were also observed (Fig. 4h and Extended Data Fig. 4e). Consistent with these findings, nuclear β-catenin immunoreactivity increased in CC nude mouse xenografts expressing the miR-100/125b bicistron and decreased in their CC-CR counterparts expressing the bicistronic sponge (Extended Data Fig. 7b).

Based on our findings that Wnt signaling is increased in CC-CR, we hypothesized that cetuximab responsiveness may be restored by suppression of Wnt signaling. Since DKK1 and DKK3 are secreted Wnt antagonists and among the most downregulated genes in CC-CR, we tested whether their overexpression could overcome cetuximab resistance using a doxycycline-inducible lentiviral system[22]. Although induction of DKK1 or DKK3 resulted in a slight reduction in colony number, this effect was augmented with addition of cetuximab (Extended Data Fig. 8a and b). Moreover, administration of recombinant DKK1 and DKK3 enhanced the ability of cetuximab to decrease proliferation and increase apoptosis (Extended Data Fig. 8c). Furthermore, nuclear β-catenin expression decreased when DKK1 or DKK3 was inducibly expressed in CC-CR in the presence of cetuximab (Extended Data Fig. 8d). We next tested whether pharmacological inhibition of Wnt activity sensitized CC-CR to cetuximab using a tankyrase inhibitor, XAV-939[23], and a β-catenin/CBP inhibitor, ICG-001[24]. Both compounds caused a concentration-dependent reduction in colony number, and cetuximab growth inhibition was enhanced by their addition (Extended Data Fig. 8e). ICG-001 also enhanced the growth inhibitory effects of cetuximab in other CRC and HNSCC cell lines with high expression of MIR100HG (Extended Data Fig. 8f and 4e). In CC-CR nude mouse xenografts, administration of cetuximab and ICG-001 individually only slowed tumor growth; however, combined treatment resulted in tumor regression (Fig. 4i-k; Extended Data Fig. 8g and h). Thus, blockade of Wnt signaling, either upstream or downstream of the APC/β-catenin degradation complex, restores cetuximab responsiveness to cetuximab-resistant cells.

Reciprocal negative regulation between GATA6 and MIR100HG/miR-125b

To explore mechanism(s) by which miR-100 and miR-125b are upregulated in CC-CR, we investigated transcriptional regulation of the host gene, MIR100HG. Possible transcription factors containing binding sites within the 2.5 kb promoter of MIR100HG were mapped *in silico* using the Match program (version 1.0)[25] and cross-referenced with the RNA-Seq

dataset (Extended Data Table 6). Among these transcription factors, we focused on the zinc-finger transcription factor GATA6, which was downregulated at both the mRNA and protein level in CC-CR in 3D culture and in nude mouse xenografts (Fig. 2a and Fig. 5a-c).

GATA6 is critical for gut endoderm development, and it both promotes and suppresses gastrointestinal and pancreatic neoplasia[26-29]. We found that MIR100HG expression decreased in cetuximab-treated CC, while GATA6 mRNA progressively increased over 48 h (Fig. 5d); however, this phenomenon did not occur in CC-CR (data not shown). GATA6 knockdown in CC (Extended Data Fig. 9a and Fig. 5e, top) caused MIR100HG upregulation and its expression no longer decreased upon cetuximab treatment (Fig. 5e, bottom), suggesting a repressive effect of GATA6 on MIR100HG. Luciferase reporter assays showed overexpression of GATA6 (Extended Data Fig. 9b) resulted in a concentration-dependent inhibition of MIR100HG promoter activity (Fig. 5f). Four putative GATA binding sites (G/A)GATA(A/T) were identified in the MIR100HG promoter region (Fig. 5g). Sequential deletions and mutations of these binding sites revealed that GATA binding site 2 (-1198 upstream of the TSS) is the major site for GATA6 repression of MIR100HG transcriptional activity (Fig. 5h). GATA6 repression of MIR100HG was also validated in HuTu80 cells with low expression of GATA6 and high expression of MIR100HG (Extended Data Fig. 9c and d). Chromatin occupancy of GATA6 at GATA-binding site 2 was confirmed by chromatin immunoprecipitation (ChIP) and electromobility shift assay (EMSA) using nuclear extracts from CC cells (Extended Data Fig. 9e and f).

Of interest, we found that the 3′ UTR of GATA6 harbors a putative binding site for miR-125b (Extended Data Table 5). In both CC and Caco-2 cells, introduction of miR-125b reduced luciferase activity of the wild-type 3′ UTR reporter construct, but not when the miR-125b site was mutated (Fig. 5i and Extended Data Fig. 9g). As predicted, GATA6 levels were reduced in CC and Caco-2 cells stably expressing miR-125b, and conversely increased in CC-CR and HuTu80 cells expressing the miR-125b sponge (Fig. 5j and Extended Data Fig. 9h). Further analysis of the TCGA data repository indicated that GATA6 is significantly downregulated, whereas MIR100HG is significantly upregulated in stage IV CRC patients (Fig. 5k). Also, CRCs with lower quartile expression of GATA6 tend to have higher expression of MIR100HG in the TCGA data repository (Fig. 5k), as well as in two additional CRC datasets (Extended Data Fig. 9i). Together, these findings suggest a double-negative regulatory circuit between GATA6 and MIR100HG/miR-125b underlies cetuximab resistance.

Increased MIR100HG, miR-100, and miR-125b are found in CRC specimens at time of progression on cetuximab

To examine whether this mode of cetuximab resistance occurs in human CRC, we obtained paired tumor specimens from ten individuals prior to the start of cetuximab treatment and at the time of tumor progression (Extended Data Table 7). *KRAS/NRAS/BRAF* mutations had been excluded in tumor specimens obtained prior to treatment with cetuximab. qRT-PCR showed that miR-100 and miR-125b were coordinately overexpressed (r_S=0.842, P<0.01) in tumors that progressed on treatment compared to pre-treatment levels (P<0.05, Fig. 6a). In addition, nuclear β-catenin immunoreactivity was significantly higher in tumors that

progressed on cetuximab (Fig. 6b). miR-125b expression directly correlated with nuclear β-catenin staining (r_s=0.636, P<0.05); the correlation between miR-100 expression and nuclear β-catenin staining did not reach statistical significance (r_s=0.612, P=0.06). Conversely and consistent with our pre-clinical findings, there was reduced nuclear GATA6 expression in tumors that advanced on cetuximab (Fig. 6c). However, we did not find a significant inverse correlation between miR-100 and GATA6 (r_s=-0.455, P=0.187), or miR-125b and GATA6 (r_s=-0.515, P=0.128). By FISH analysis, the MIR100HG, miR-100, and miR-125b signals increased in tumors that progressed on treatment. In these same samples, there was increased β-catenin staining and reduced GATA6 staining (Fig. 6d). We excluded *MET* amplification by FISH in all ten paired specimens and sequenced the post-treatment tumors for mutations in *KRAS/NRAS/BRAF* (Extended Data Fig. 9j and Extended Data Table 8). *NRAS* and *KRAS* mutations were detected in 2 cases, respectively; we confirmed that these were likely acquired events by re-sequencing the pre-treatment DNA. In both cases, MIR100HG and miR-100/125b were increased. In the remaining 8 cases that lacked genetic resistance events, 5 cases exhibited upregulated MIR100HG and miR-100/125b in the tumors post-treatment. These clinical data support our pre-clinical findings and demonstrate that upregulation of MIR100HG and miR-100/125b occur in the setting of acquired cetuximab resistance in CRC patients, and this upregulation may both coincide with and be independent of genetic mutations associated with cetuximab resistance.

Discussion

MIR100HG is a polycistronic miRNA host gene, which encodes miR-100, let-7a-2, and miR-125b-1 within its third intron. MIR100HG was first reported to participate in fate determination of human mesenchymal stem cells[30], and later found to be highly expressed in acute megakaryoblastic leukemia[12,13]. Increased MIR100HG expression is associated with a poor prognosis in cervical cancer[31], whereas its expression is reduced in breast cancer due to hypermethylation[32]. Increased expression of miR-100 and miR-125b are also correlated with gastric cancer progression in clinical samples[33]. In our study, concomitant upregulation of MIR100HG and miR-100/125b occurs in the setting of acquired and *de novo* cetuximab resistance in CRC and HNSCC cell lines. Moreover, we show these events can co-occur with *KRAS*/*NRAS/BRAF* mutation and in tumors of CRC patients that progressed on cetuximab. Analysis of TCGA CRC data repository revealed a stage-dependent increase of MIR100HG expression. These data support the hypothesis that MIR100HG and miR-100/125b are potential predictive biomarkers for cetuximab resistance.

We identified that miR-100 and miR-125b coordinately contribute to cetuximab resistance by targeting five negative regulators of Wnt signaling. miR-100 targets DKK1 and ZNRF3; miR-125b also targets ZNRF3, as well as RNF43, DKK3, and APC2. Wnt signaling is tightly regulated and negative regulators act at many different levels[34,35]. DKK1 and DKK3 are secreted Wnt signaling antagonists of the Dickkopf family. DKK1 acts by binding and internalizing the Wnt co-receptor LRP5/6[36], while it is unclear how DKK3 attenuates Wnt signaling[37]. ZNRF3 and RNF43 are two closely related transmembrane E3 ubiquitin ligases that antagonize Wnt signaling through ubiquitylation and degradation of the Wnt receptor Frizzled and its co-receptor LRP5/6[38,39]. Although inactivating mutations have been reported for *ZNRF3/RNF43*[40], our data suggest that downregulation by miR-100/125b may

represent an alternative mechanism of attenuating ZNRF3/RNF43 function. APC2 targets β-catenin for destruction and is functionally complementary to APC[41]; it was recently reported that APC2 recruits TNKS into the β-catenin destruction complex to regulate β-catenin proteolysis[42]. We previously identified APC as a target of miR-125b in leukemia cells[12], but did not examine APC in CC or CC-CR since it is mutated in these cells. However, it was recently reported that miR-125b targets APC in mutant β-catenin HCT116 cells that have wild-type *APC*[43]. The present study shows that miR-100 and miR-125b work together to target these five Wnt negative regulators, providing a novel regulatory mode for clustered miRNAs to cooperatively regulate this pathway.

We cannot exclude that miR-100/125b contributes to cetuximab resistance through means other than Wnt signaling. For example, miR-125b can enhance tumor formation in the skin by targeting vacuolar protein-sorting 4 homolog B (Vps4b) and indirectly prolonging EGFR activity[44]. However, we observed no differences in VPS4b expression between CC and CC-CR. We also have not excluded an effect of the full-length 3 kb MIR100HG transcript on Wnt signaling.

The present study adds to the literature describing crosstalk between EGFR and Wnt signaling[45,46]. For example, in *APC-mutant* CRC, increased EGFR signaling enhances Wnt activity, supporting the notion that Wnt signaling is further modulated in the setting of an impaired β-catenin degradation complex[47]. In a reciprocal manner, binding of Wnt ligands to their GPCR Frizzled receptors results in EGFR transactivation via metalloprotease-dependent, cell-surface ectodomain cleavage of EGFR ligands[48]. Moreover, increased Wnt signaling confers resistance to EGFR tyrosine kinase inhibitors in lung cancer[49,50].

The precise mechanism by which increased Wnt signaling confers cetuximab resistance is uncertain. It has been reported that Wnt signaling increases EGFR expression in liver[51]. We observed that cetuximab did not reduce p-EGFR, pERK1/2, or p-AKT in CC-CR as it does in CC. Although there is equivalent cell-surface EGFR staining in CC and CC-CR, this does not exclude differences in rates of EGFR internalization, recycling and degradation. Going forward, a system-wide approach should prove useful to help unravel mechanisms underlying the EGFR/Wnt crosstalk in this system.

The role of GATA6 in cancer is complex and context-dependent; even in the same tumor type, conflicting evidence exists. For example, GATA6 promotes pancreatic carcinogenesis by activating Wnt signaling[52]. In separate studies, it serves a tumor suppressive role by maintaining a pancreatic differentiation program[53,54]. Likewise, in colonic neoplasia, opposing actions are reported. In colonic adenomas, GATA6 represses BMP expression, thereby enabling stem cell self-renewal[27], and in CRC cell lines, it enhances expression of Lgr5 and REG4 to promote clonogenicity and growth, respectively[55,56]. In contrast, our study supports a tumor-suppressive role for GATA6 in CRC. Analysis of the TCGA CRC repository reveals reduction in GATA6 expression in stage IV CRC along with increased MIR100HG expression. Reduced expression of GATA6 would permit increased expression of MIR100HG, and the corresponding increased expression of miR-125b would reinforce repression of GATA6. In this context, GATA6 serves a permissive tumor suppressive role by preventing Wnt signaling-enhanced cetuximab resistance.

The present study has important therapeutic implications for CRC and HNSCC. It is increasingly appreciated that there are gradients of Wnt signaling in CRC[57] and that Wnt signaling can be modulated in the setting of APC loss-of-function[58,59]. In our model, MIR100HG- and miR-100/125b-mediated Wnt activation represents an adaption of activating compensatory pathways for cells to survive under EGFR inhibition. We show induction of DKK1 or DKK3 individually, or the combined addition of recombinant DKK1 and DKK3, overcomes cetuximab resistance in CC-CR. Both XAV-939, a tankyrase inhibitor, and ICG-001, a β-catenin-CBP inhibitor, augment the growth inhibitory effects of cetuximab. We propose that future trials in individuals with wild-type *KRAS/NRAS/BRAF* CRC should consider the levels of MIR100HG expression.

In summary, we have identified a complex circuitry underlying cetuximab resistance by upregulation of MIR100HG and its embedded miRNAs (see Extended Data Fig. 10). miR-100 and miR-125b coordinately activate Wnt signaling by reducing expression of five negative regulators of Wnt signaling. miR-125b reinforces upregulation of MIR100HG by inhibiting GATA6 expression, which normally suppresses MIR100HG. We show that inhibition of Wnt signaling can overcome this mode of cetuximab resistance, underscoring the potential clinical relevance of the interactions between EGFR and Wnt signaling.

Data availability

RNA-Seq and small RNA-Seq data are available at the NCBI Gene Expression Omnibus (GEO) repository with accession code GSE82236. Whole-exome sequencing data are available at the GEO repository with accession code GSE76352. A Life Sciences Reporting Summary is available online.

Online Methods

2D and 3D cell culture

NCI-H508, Caco-2, SW403, SW948, HT29, SK-CO-1, DLD-1, SW480, SW837, SW48, SW620, LoVo, COLO205, T84, LS174T, NCI-H716, HCT8, HCT15, SW1116, RKO, COLO320DM, HuTu80, LS123, and HCT116 cell lines were from the American Type Culture Collection (ATCC). HCA-7, its derivatives CC and CC-CR, DiFi, GEO, LIM1215, and LIM2405 were maintained in the Coffey lab. The SNUC4 cell line was from the Korean Cell Line Bank and the V9P cell line was provided by John Mariadason (Olivia Newton-John Cancer Research Institute, Melbourne, Australia). HNSCC cell lines SCC25, its derived cetuximab-resistant sublines (CTX-R1, R3, R4, R5, R7, and R8), and UNC10 were maintained in Christine Chung's laboratory. All cell lines were confirmed to be free of mycoplasma contamination. Cells were grown in Dulbecco's Modified Eagle's Medium (DMEM, Corning) supplemented with 10% bovine growth serum, glutamine, nonessential amino acids, 100 U/ml penicillin and 100 μg/ml streptomycin (HyClone) in 5% CO_2 at 37°C. 3D collagen cultures were set up using 3 layers of type-I collagen PureCol (Advanced BioMatrix) in triplicate as previously described[14,15]. Human recombinant DKK1 (rDKK1) and DKK3 (rDKK3) are from R&D Systems. Drugs are used as follows: cetuximab (Merck KGaA), Wnt pathway inhibitor ICG-001 and XAV-939 (Selleck Chemicals). Colonies were counted by GelCount colony counter (Oxford Optronix).

RNA-Seq analysis

Total RNA from cells embedded in collagen was isolated by TRIzol Reagent (Invitrogen) and then purified using RNeasy Mini Kit (Qiagen). The concentration and integrity of total RNA were estimated using the Qubit 2.0 Fluorometer (Invitrogen) and Agilent 2100 Bioanalyzer (Agilent Technologies), respectively. Polyadenylated RNAs were isolated using NEBNext Magnetic Oligo d(T)25 Beads. First strand synthesis was performed using NEBNext RNA First Strand Synthesis Module (New England BioLabs). Directional second strand synthesis was performed using NEBNext Ultra Directional Second Strand Synthesis Module. The NEBNext DNA Library Prep Master Mix Set for Illumina was used to prepare next-generation sequencing expression libraries per manufacturer's protocol. Accurate quantification for sequencing applications was determined using the qPCR-based KAPA Biosystems Library Quantification Kit (Kapa Biosystems). Paired-end (PE) sequencing (75bp) was performed on the NextSeq 500 sequencer (Illumina).

RNA-Seq reads were aligned to the human genome hg19 using TopHat2[60], and the number of reads mapped to each gene was calculated by HTseq (http://www-huber.embl.de/users/anders/HTSeq/). Differentially expressed genes between CC and CC-CR were detected by edgeR based on negative binomial distribution[61]. The *p*-values were adjusted by Benjamini and Hochberg's multiple test correction procedures. Differential expression was determined based on fold-change (FC) and false discovery rate (FDR) with $|\log_2(FC)| > 1$ and FDR<0.01.

Small RNA-Seq analysis

Approximately 1 μg of total RNA from each sample was utilized for small RNA library preparation using NEBNext Small RNA Library Prep Set for Illumina (New England BioLabs) following the manufacturer's protocol. Post PCR material was purified using QIAquick PCR Purification Kit (Qiagen). Post PCR yield and concentration of the prepared libraries were assessed using Qubit 2.0 Fluorometer and DNA 1000 chip on Agilent 2100 Bioanalyzer. Size selection of small RNA was done on the Pippin Prep instrument (Sage Science). Accurate quantification for sequencing applications was performed using the qPCR-based KAPA Biosystems Library Quantification Kit. Single end sequencing (50BP) was performed on a NextSeq 500 Sequencer (Illumina).

Adapters from 3′ end of small RNA-Seq reads were trimmed by Cutadpt (http://code.google.com/p/cutadapt/). Sequences shorter than 15 bp were excluded from the downstream analysis. Reads were aligned to the human genome hg19 using Bowtie. Mapped reads were annotated, and miRNA expression was quantified using ncPRO-seq (version v1.5.1)[62] based on miRbase v19. Differentially expressed miRNAs between CC-CR and CC cells were detected by edgeR[61]. The *p*-values were adjusted by Benjamini and Hochberg's multiple test correction procedures. Differential expression was determined based on FC and FDR with $|\log_2(FC)| > 1$ and FDR <0.01.

Whole-exome sequencing

DNA extraction was performed by the QIAamp DNA Mini Kit (Qiagen) following the manufacturer's instructions. The DNA was quantified by Nanodrop spectrophotometer

(Thermo Fisher Scientific). Genomic DNA was sequenced using Illumina HiSeq 2500. Reads were aligned to the human genome hg19 with BWA, sorted, and indexed with SAMtools. Duplicated reads were marked by Picard (http://picard.sourceforge.net/). SNPs and Indels were called simultaneously on CC and CC-CR samples by SAMtools with base quality ≥30, reads with mapping quality ≥30, and mapping quality downgrading coefficient of 50. SNPs and Indels with strand bias $P<0.01$, base quality bias $P<0.01$, mapping quality bias $P<0.01$, or end distance bias $P<0.01$ were filtered out. Furthermore, SNPs within 3 bp around a gap were removed. SNPs and Indels were annotated, and their effects were predicted by snpEff and snpSift[63]. Strekla[64] was used to detect SNVs and indels that were present at a significantly different frequency between CC and CC-CR samples with default parameters except turning off the depth filter for exome sequencing data.

Constructs, oligonucleotides, infection and transfection of human cell lines

miRNA expression lentiviral vectors LeGO-cO:miR-100, LeGO-cO:miR-125b, LeGO-cO:miR-100/125b bicistron, and control empty vector were used as described previously[12]. Lentivirus produced in HEK293 cells was generated and collected using standard protocols[12]. GFP-positive infected cells were selected in Blasticidin S (10 μg/ml) followed by flow sorting. Stable miRNA knockdown was achieved by introducing the lentiviral miRNA sponge constructs that target either miR-100, miR-125b alone or both. Briefly, eight repeats of anti-sense miR-100 (5′-CACAAGTTCGGATCTACGGGTT-3′) or/and anti-sense miR-125b (5′-TCACAAGTTAGGGTCTCAGGGA-3′) were designed and synthesized following standard protocols[65], and then cloned into the pGLV3/H1/GFP vector (GenePharma). A control sponge was used, which includes eight repeats of an artificial miRNA (5′-AAGTTTTCAGAAAGCTAACA-3′)[66] that is not complementary to any known miRNA. GFP-positive infected cells were selected in puromycin (1 μg/ml) followed by flow sorting.

Human DKK1 expression vector pcDNA3-DKK1-FLAG was kindly provided by Dr. Stuart Aaronson (Mount Sinai School of Medicine). Human DKK3 expression vector pCS2-DKK3-flag was from Addgene (plasmid #15496). Lentiviral-inducible expression constructs containing DKK1 or DKK3 under the control of a doxycycline-inducible promoter were constructed by transferring each ORF into the pInducer-20 lentiviral vector[67]. Infected cells were selected in G418 (200 μg/ml) to generate stable cell lines. Tetracycline-reduced FBS (Clontech) was substituted for all media for cells transduced with the pInducer-20 vectors. To induce expression of DKK1 or DKK3, 1 μg/ml doxycycline (Sigma-Aldrich) was added to the culture medium.

pcDNA3.1-GATA6 and pcDNA3.1-mutant (mut) GATA6 by site-directed mutagenesis were kindly provided by Dr. Christine A. Iacobuzio-Donahue (Memorial Sloan Kettering Cancer Center). GATA6 Silencer Select siRNAs (ID s5605, s5606) and the Silencer Select negative control siRNA were used for transient transfection (Life Technologies). GATA6 expression plasmids or siRNAs were transfected into indicated cells using Lipofectamine 2000 or RNAiMAX Reagent (Thermo Fisher Scientific), respectively. Experiments were performed 48 h after transfection.

For 3′ UTR luciferase reporter assay, the 3′ UTR fragments of DKK1, DKK3, ZNRF3, RNF43, and APC2 containing miR-100 or miR-125b putative target sites were amplified and cloned downstream of the SV40 promoter-driven *Renilla* luciferase cassette in psiCHECK-2 (Promega). For luciferase reporter assays to measure promoter activities, PCR products of sequential deletion fragments of human MIR100HG promoter were cloned into pGL3-Basic vector (Promega). A site-directed mutagenesis kit (Agilent Technologies) was used to mutate the miR-100, miR-125b, or GATA6 binding sites of these vectors. All sequences were confirmed by sequencing.

Quantitative RT-PCR

Analysis of mRNA and miRNA levels was performed on the StepOnePlus Real-Time PCR System (Applied Biosystems). For mRNA detection, cDNA was generated with the QuantiTect Reverse Transcription Kit (Qiagen). Diluted cDNA samples were amplified to establish a standard curve for calculation of relative target concentrations using Express SYBR GreenER qPCR SuperMix with Premixed ROX (Life Technologies). The housekeeping gene ACTB was used as an internal control. The primers for the genes of interest were synthesized by RealTimePrimers.com or Sigma-Aldrich (Extended Data Table 9). Analysis of lncRNA and miRNA levels was performed with the use of the TaqMan fast advanced master mix (Applied Biosystems). TaqMan lncRNA, miRNA, and Pri-miRNA expression assays (Life technologies) were used according to the manufacturer's instructions, with ACTB or U6 small nuclear RNA (U6 snRNA) as the internal control (Extended Data Table 10). The relative expression of RNAs was calculated using the comparative Ct method.

5′ Rapid amplification of cDNA ends (RACE)

5′ RACE was used to determine transcriptional initiation sites of lncRNA MIR100HG using FirstChoice RLM-RACE Kit (Thermo Fisher Scientific), and Zero Blunt TOPO PCR Cloning Kit (Life Technologies) was used for sequencing according to the manufacturer's instructions. Two reverse primers for the TSS of MIR100HG were used in a nested PCR with the two 5′ primers from the kit. Outer primer: 5′-AAACCGGGCCCTCCAGTTCACTAT C-3′; Inner primer: 5′-TCTTTTCCATCCCCTTTGCATGTGG-3.

Western blot analysis

Whole cell lysates were prepared using RIPA buffer supplemented with protease inhibitor cocktail (Sigma-Aldrich) and phosphatase inhibitor (Roche). The nuclear extract was isolated using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Fisher Scientific). Protein concentration was determined using the BCA Protein Assay Kit (Thermo Fisher Scientific). Primary antibodies were against DKK1 (Santa Cruz sc-25516), ZNRF3 (sc-86958), RNF43 (sc-165398), GATA6 (sc-7244), Lamin A/C (sc-7292), and GAPDH (sc-20357); DKK3 (Abcam #2459), APC2 (Abcam #80018), ZNRF3 (Abcam #122353), RNF43 (Abcam #84125), and p-EGFR Y1068 (Abcam #5644); β-Catenin (BD Biosciences #610154); EGFR (Millipore #06-847) and p-LRP6 (Millipore #07-2187); ERK1/2 (Cell Signaling Technology, CST, #4695), p-ERK1/2 (CST #4370), AKT (CST #9272), p-AKT (CST #4060), cleaved Caspase-3 (CST #9664), Cyclin-D1 (CST #2978), BIM (CST #2933),

p-β-catenin (S552) (CST #9566), LRP6 (CST #2560), and GATA6 (CST #5851); β-actin (Sigma-Aldrich A1978). Immunoreactivity was detected and the signals were analyzed under nonsaturating conditions with an image densitometer (ChemiDoc MP Imager, Bio-Rad) and Image Lab software (Bio-Rad). All immunoblot analyses were performed at least three times.

Immunofluorescent staining

Immunofluorescent staining and confocal analysis were performed as described[14]. Primary antibodies were against cleaved Caspase-3 antibody (Abcam #2302), Ki-67 (Dako M7240), p-β-catenin (Y489) (Developmental Studies Hybridoma Bank, IA), and GATA6 (CST #5851). Alexa Fluor 568 Phalloidin was from Life Technologies. Secondary antibodies included Alexa Fluor 488, 568, or 647-conjugated goat anti-mouse or anti-rabbit IgGs (Life Technologies). Nuclei were stained with Hoechst 33342. Slides were mounted with Prolong Gold Antifade Reagent (Life Technologies) prior to imaging on a Zeiss LSM 710 confocal microscope. Quantification was done by manually counting at least 5 randomly chosen high power fields (HPFs) per sample.

Luciferase reporter assays

For 3′ UTR luciferase reporter assays, indicated cells cultured in 24-well plates were co-transfected with miR-100 or miR-125b mimic or negative control (Ambion) and indicated psiCHECK-2-3′ UTR wild-type or mutant plasmids using Lipofectamine 2000 (Thermo Fisher Scientific). *Renilla* and firefly luciferase activities were measured after 48 h with the dual-luciferase reporter assay system (Promega). *Renilla* luciferase activity was normalized to firefly activity and presented as relative luciferase activity.

For luciferase reporter assay to measure promoter activities, indicated cells were co-transfected with pGL3-MIR100HG promoter fragment, pRL-SV40 *Renilla* luciferase reporter, and pcDNA3.1-GATA6 or mut GATA6 expression plasmid or empty vector control. The firefly and *Renilla* luciferase activity was measured with the dual-luciferase reporter assay system (Promega). Firefly luciferase activity was normalized to *Renilla* activity and presented as relative luciferase activity. All assays were performed in triplicate three times.

Chromatin immunoprecipitation (ChIP)

ChIP assays were performed using a Pierce Agarose ChIP Kit (Thermo Fisher Scientific) following the manufacturer's instructions, and ChIP-enriched DNA samples were analyzed by qPCR. Cells were cross-linked with 1% formaldehyde for 10 min at room temperature and quenched in glycine. Rabbit anti-GATA6 antibody (CST #5851) or normal rabbit IgG (BD Biosciences) were used for immunoprecipitation. The DNA was recovered and subjected to qPCR to amplify the binding sites of the MIR100HG promoter region. Data are presented as relative enrichment normalized to control IgG. The qPCR primer sets are: for GATA-binding site-2 (-1198 upstream of the TSS): forward, 5′-ACCTATCTCTGCTACTTATTTTATG-3′, reverse, 5′-CTATTTATCAGCACAGTTACTGG-3′; distal region primer sets: forward, 5′-GAATGCAGTAGTGGCTAG GAATG-3′, reverse, 5′-CTAACTCTCTAGGCTGTTATCTG-3′.

Electrophoretic mobility shift assay (EMSA)

Nuclear and cytoplasmic proteins were extracted using the NE-PER Nuclear and Cytoplasmic Extraction Reagents Kit (Thermo Fisher Scientific). LightShift Chemiluminescent EMSA Kit (Thermo Fisher Scientific) was used according to the manufacturer's protocol. Double-stranded biotin-labeled probe (5′-GCTCTCTATTTATCA GCACAGTTA-biotin-3′) was used, which corresponds to GATA-binding site-2 of the human MIR100HG gene promoter (-1198 upstream of the TSS). Nuclear extracts (8 μg) were incubated with labeled probe, poly (dI-dC), and the binding buffer for 30 min at room temperature. For supershift experiments, 1 μg of monoclonal rabbit anti-GATA6 (CST #5851) was added to the reaction mixture before the addition of labeled probe. For the binding competition experiment, an excess (200-fold) of unlabeled cold competitor probe or mutant probe (5′-GCTCTCTGCCCGCTGATACAGTTA-3′) was added into the reaction mixture. Bound DNA complexes were resolved on polyacrylamide gels and transferred to a nylon membrane (Thermo Fisher Scientific). Nylon membranes were cross-linked and chemiluminescent detection was performed.

Immunohistochemistry (IHC)

IHC for target molecules was performed on serial sections from tumor tissues of nude mice xenografts and CRC patients. Tissue sections were deparaffinized, subjected to antigen retrieval using target antigen retrieval solution (Dako), and incubated with primary antibodies against Ki67 (Dako M7240), cleaved Caspase-3 (CST #9661), β-catenin (BD Biosciences #610154), GATA6 (Abcam #22600), DKK1 (Abcam #61034), DKK3 (Abcam #115869), ZNRF3 (Abcam #122353), RNF43 (Abcam #129401), and APC2 (Abcam #113370). Then sections were incubated with Envision System HRP-labeled polymer anti-rabbit or anti-mouse secondary antibodies (Dako). The results of IHC were scored by two independent observers. Ki-67 and cleaved Caspase-3 staining was quantified by calculating positively stained cells in at least 5 randomly chosen HPFs of each sample. Quantification of other molecules was based on intensity and extent of staining according to the histological scoring method as previously described[68].

In vivo tumor growth in xenograft model

In vivo cetuximab treatment was performed using 6~8-week-old female athymic BALB/c nude mice. All experiments were conducted under protocols approved by the Fourth Military Medical University Institutional Animal Care and Use Committee. Suspensions of the corresponding cells were subcutaneously injected into the flanks (6×10^6 tumor cells/150 μl PBS per spot; 6-8 mice in each group). Animals were weighed, and the tumor size was measured using bilateral caliper measurements. Tumor volume was calculated using the formula: Tumor maximum diameter (L) × the right angle diameter to that axis $(W)^2/2$. When the tumors reached the determined size (around 100 mm^3), mice were randomized into control and treatment groups. Cetuximab treatment was given at a dose of 1 mg/mouse, intraperitoneal (i.p.) injection, every 3 days. For Wnt pathway inhibitor ICG-001 *in vivo* treatment, the sodium phosphate form of ICG-001, synthesized by Vanderbilt Institute of Chemical Biology (VICB) Synthesis Core, was administered at a dose of 150 mg/kg body weight, i.p. injection every day. *In vivo* imaging system (IVIS, PerkinElmer) was used to

detect GFP fluorescence in tumor-bearing mice. After 4 weeks of treatment, mice were sacrificed according to institutional ethical guidelines. Postmortem examination included tumor size and weight measurements, and then tumors were paraffin-embedded to perform hematoxylin and eosin (H&E) staining. The sample size for the experiments was based on the pilot studies and determined to ensure a power at 0.8 with type 1 error (α) at 0.05 of expected difference. Postmortem examination and data analysis were done by two investigators blinded to the group allocations.

Human CRC samples and subjects

All human CRC samples were obtained from the Xijing Hospital of Digestive Diseases. The study was approved by the Ethics Committee of Xijing Hospital with written informed consent obtained from all subjects. The pathological status of the specimens was provided by the Department of Pathology. In total, we analyzed 10 pairs of tumor specimen pre- and post-cetuximab treatment. Pre-cetuximab treated specimens were retrospectively obtained during surgical or biopsy under colonoscopy on subjects with CRC. After computed tomography (CT) of tumor lesions demonstrated tumor re-growth (disease progression) following initial response to cetuximab-based therapy, post-cetuximab treated specimens were collected whenever possible at the time of progression. RNA was extracted with RecoverAll Total Nucleic Acid Isolation Kit for FFPE (Ambion). IHC and *FISH* analyses were performed on formalin-fixed, paraffin-embedded (FFPE) tumor tissue *sections.* Blind evaluation was done by two pathologists.

Fluorescence *in situ* hybridization (FISH) assays

Locked nucleic acid-*in situ* hybridization (LNA-ISH) with tyramide signal amplification (TSA) was performed to detect lncRNA and miRNA as previously described[13,69]. All LNA probes were synthesized (Exiqon) including double biotin-labeled probe against MIR100HG, double digoxigenin (DIG)-labeled probe against miR-125b, double fluorescein-labeled probe against miR-100, DIG-labeled probe against U6 snRNA, and DIG-labeled scramble probe. Anti-Digoxigenin HRP Conjugate, anti-Fluorescein HRP Conjugate, Streptavidin-HRP Conjugate, and TSA Cy3 and Fluorescein Kit (all from PerkinElmer) were used for TSA methods. Confocal fluorescence microscopy was performed using a Zeiss LSM 710 confocal microscope.

For detection of *MET* amplification, *MET*/*CEP7* dual-color probes (Cytotest) were used for recognizing the *MET* gene status following the manufacturer's protocol. Analysis was according to the University of Colorado Cancer Center (UCCC) criteria. A *MET*/*CEP7* ratio was established based on counting at least 200 cells.

Targeted Sanger sequencing of *KRAS, NRAS* and *BRAF*

KRAS/*NRAS*/*BRAF* mutations had been excluded in tumor specimens obtained prior to treatment with cetuximab. This study was carried out on 12 FFPE blocks of colorectal carcinomas (10 blocks obtained after disease progression upon cetuximab treatment and case 2 and 4 blocks before treatment). Genomic DNA was isolated using QIAamp DNA FFPE Tissue Kit (QIAGEN). Oncogenic alleles of *KRAS* (codon G12, G13, Q61, K117, and A146), *NRAS* (codon G12, G13, and Q61), and *BRAF* (codon G465, G468, Y472, D593,

F594, L596, L597, T598, V600, and K601) were sequenced by targeted Sanger sequencing with PCR primers listed (Extended Data Table 11). The PCR products were then sequenced using BigDye Terminator 3.1 Cycle Sequencing Kit on a 96-capillary 3730XL DNA Sequencer (Applied Biosystems).

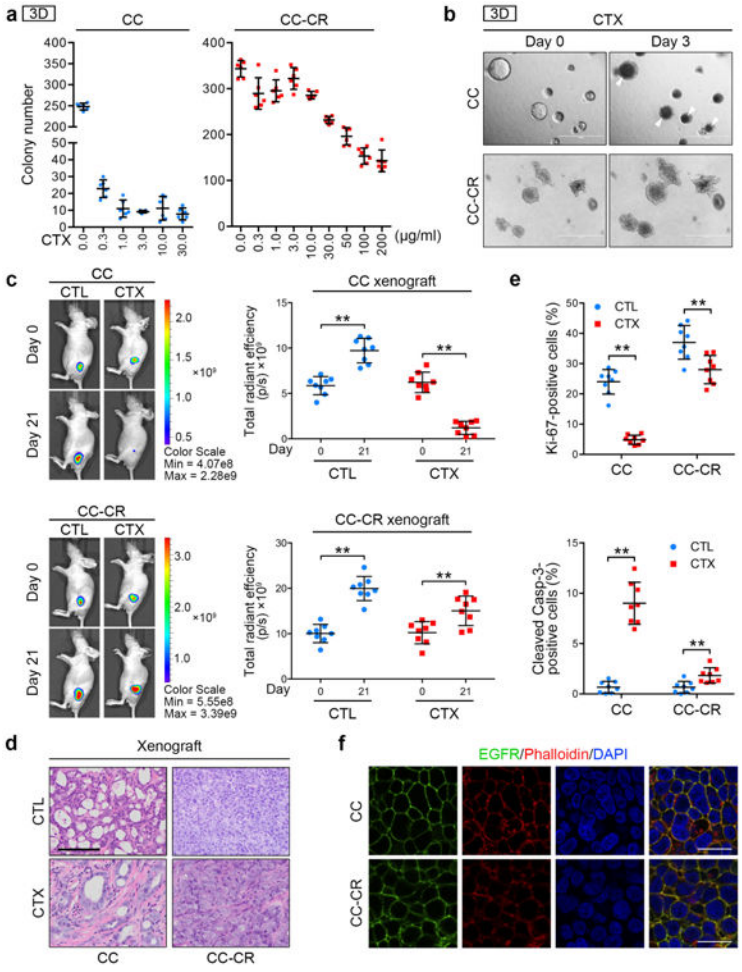
Statistical analysis

Statistical analysis was performed by the SPSS 18.0 (SPSS Inc.) and R (version 3.3.1). The statistical significance between data sets was expressed as *P*-values, and $P<0.05$ was considered statistically significant. Two-tailed unpaired or paired Student's *t* test, ANOVA (Dunnett's or LSD post-hoc test), non-parametric signed rank test, Mann– Whitney *U* test, and Pearson correlation coefficients were used according to the type of experiment.

For Pearson correlation of MIR100HG expression with 64-gene Wnt scores and RAS_AZ scores on 458 CRCs, the Log_2 expression values of MIR100HG were obtained on 458 CRC tumors (10 samples without suitable microarray data were excluded from 468 CRCs previously reported[19,70]). A set of 64 “consensus” β-catenin (upregulated) targeted genes were adopted from a recent study[71], and a mean Log_2 expression of the 64 genes was calculated as the Wnt/β-catenin pathway score on 458 CRC samples as previously described[19]. The RAS_AZ signature score, which measures MEK activation as a downstream index of RAS activity, was previously developed[20] and was pre-calculated on 458 CRC samples[19,70]. For comparison of the Wnt scores between CC and CC-CR, the 64-gene Wnt scores (Log_2 expression) were calculated for CC and CC-CR (each with 3 replicates) similarly as described above and then subjected to the two-tailed unpaired Student's *t* test.

RNA-Seq and small RNA-Seq data of CRC was obtained from TCGA Firehose developed by the Broad GDAC (https://confluence.broadinstitute.org/display/GDAC/ Dashboard-Stddata). mRNA expression array data of CRC/colon cancer was acquired from published studies (GEO accession: GSE14333, GSE39582)[72,73]. The gene expression abundances were Log_2 transformed, and Pearson correlation coefficients were used to measure the correlation between MIR100HG and miR-100, miR-125b, and let-7a. Mann-Whitney *U* test was used to determine the expression difference of MIR100HG in the lower (0-25%) and the higher (>75%) quartiles of GATA6 expression.

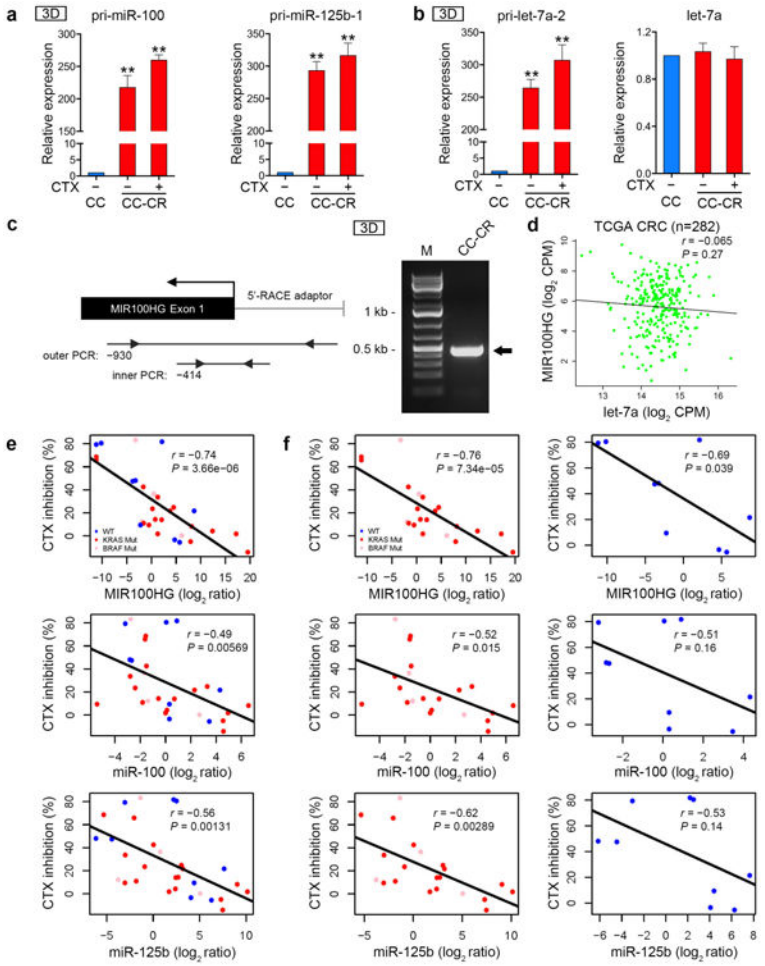
Extended Data



Extended Data Fig. 1.

Establishment of cetuximab-resistant cells in 3D culture.

(a) Five thousand cells/ml were cultured in type-1 collagen for 17 days. Fresh medium was added with different concentrations of CTX every 2 days, and colony number was determined using a GelCount plate reader. n=2 independent experiments performed in triplicate. **(b)** Twelve-day old CC and CR were treated with CTX (10 μg/ml) for 3 days. Representative images from 3 independent experiments are shown. Scale bars, 1000 μm. **(c)** Left: Representative fluorescence images of GFP signals captured from subcutaneous tumors, generated by injection of CC and CC-CR stably transduced with GFP-expressing lentivirus. Right: Quantification of radiant efficiency from tumors. n=8. $^{**}P<0.01$ by paired Student's *t* test. **(d)** Representative H&E staining of the tumor xenografts from the indicated groups. Scale bar, 100 μm. **(e)** Quantification of IHC staining in Fig. 1g. n=8 mice. $^{**}P<0.01$ by Student's *t* test. **(f)** CC and CC-CR cells grown on Transwell filters were incubated with Alexa Fluor 488-labeled C225 mAb directed against the extracellular domain of EGFR and then stained for F-actin (Phalloidin) and nuclei (DAPI). Scale bars, 20 μm. Data represent mean ± s.d. in a, c, and e.

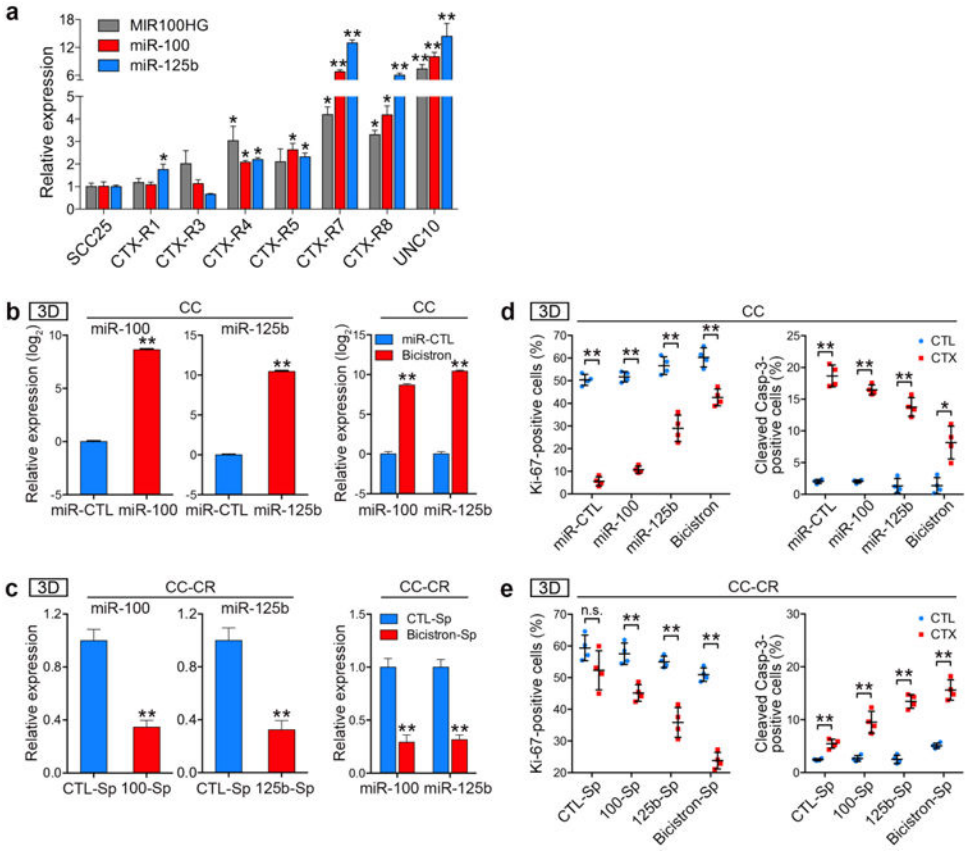


Extended Data Fig. 2.

MIR100HG and miR-100/125b overexpression in cetuximab-resistant colorectal cancer cell lines.

(a) qRT-PCR showing upregulation of pri-miR-100 and pri-miR-125b-1 in CC-CR compared to CC grown in 3D. In CC-CR, cells were treated with CTX (CTX+, 3 μg/ml) or normal culture medium (CTX-) for consecutive 14 days in 3D. **(b)** qRT-PCR showing upregulation of pri-let-7a-2 expression in CC-CR but unchanged expression of mature let-7a between the 2 cell lines. n=3 independent experiments performed in triplicate in a and b. Data represent mean ± s.d. $^{**}P<0.01$ by one-way ANOVA followed by Dunnett's test compared with CC. **(c)** Left: a schematic diagram showing the PCR primers used in the 5′ RACE. Right: MIR100HG TSS was validated by 5′ RACE nested PCR in CC-CR with subsequent sequencing of the cloned fragments. Arrow indicates band of expected size. M, DNA marker. **(d)** Scatter plots of MIR100HG versus let-7a expression in TCGA CRC data repository. No correlation was found between those 2 molecules. **(e, f)** Expression of MIR100HG and miR-100/125b negatively correlates with cetuximab growth inhibition regardless of *KRAS*/*BRAF* mutational status. (e) Scatter plot of MIR100HG and miR-100/125b expression versus cetuximab inhibition rate in a panel of 30 CRC cell lines.

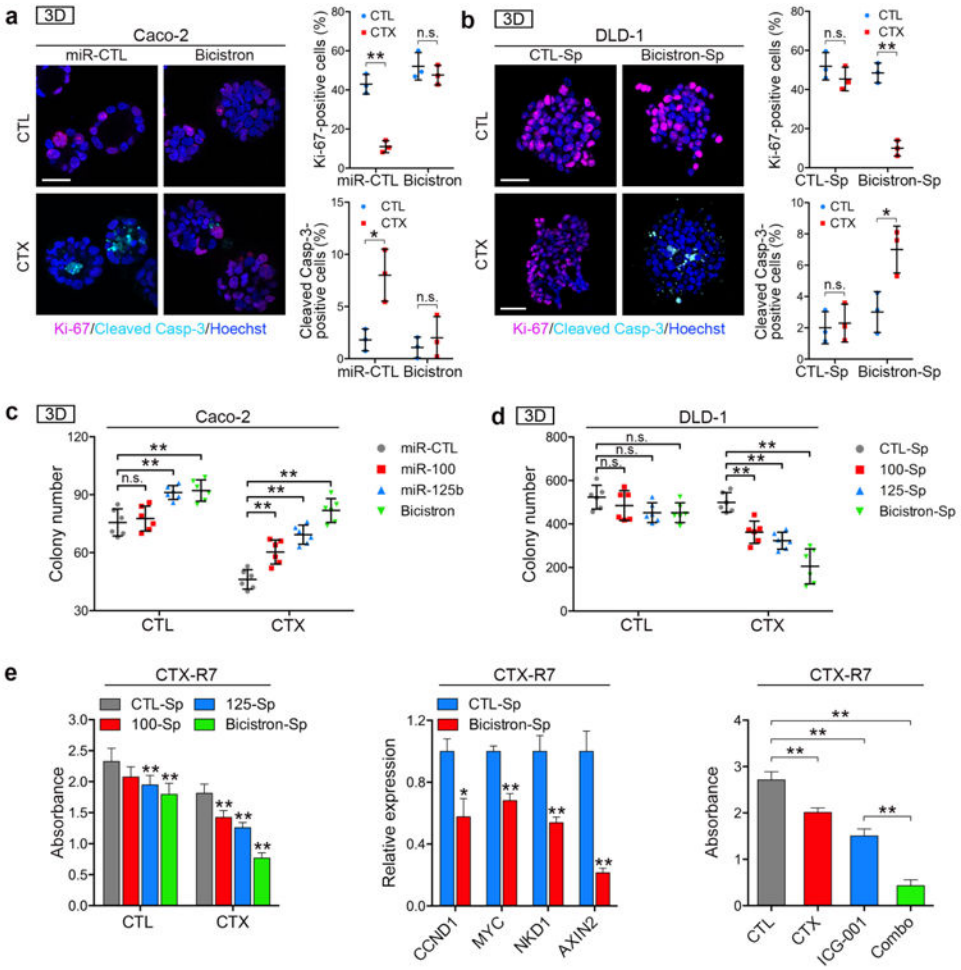
(f) Twenty-one cell lines harbor *KRAS* or *BRAF* mutation, and 9 cell lines are *KRAS*/*BRAF* wild-type (WT). Pearson correlation coefficients (*r*) and *P* values are shown.



Extended Data Fig. 3.

MIR100HG and miR-100/125b expression in head and neck squamous cell cancer cell lines and modulation of miR-100 and/or miR-125b in CC and CC-CR cells.

(a) qRT-PCR analysis of MIR100HG, miR-100, and miR-125b expression among the CTX-sensitive head and neck squamous cell carcinoma (HNSCC) cell line SCC25 and its derived CTX-resistant sublines (CTX-R1, R3, R4, R5, R7, and R8) upon continuous exposure to cetuximab, as well as UNC10, a *de novo* CTX-resistant cell line. n=3 independent experiments performed in triplicate. **P*<0.05, ***P*<0.01 by one-way ANOVA followed by Dunnett's test compared with SCC25. **(b)** qRT-PCR of indicated miRNA expression in CC stably overexpressing miR-100, miR-125b, or Bicistron. **(c)** qRT-PCR of indicated miRNA expression in CC-CR stably expressing miR-100 sponge (100-Sp), miR-125b sponge (125b-Sp), or bicistron sponge (Bicistron-Sp). Values were normalized to U6 snRNA. n=3 experiments performed in triplicate. ***P*<0.01 by Student's *t* test. **(d, e)** Quantification of Ki-67 and Cleaved Casp-3 in Fig. 3c and d. n=4 independent experiments. **P*<0.05, ***P*<0.01 by Student's *t* test. Data represent mean ± s.d. n.s., not significant.

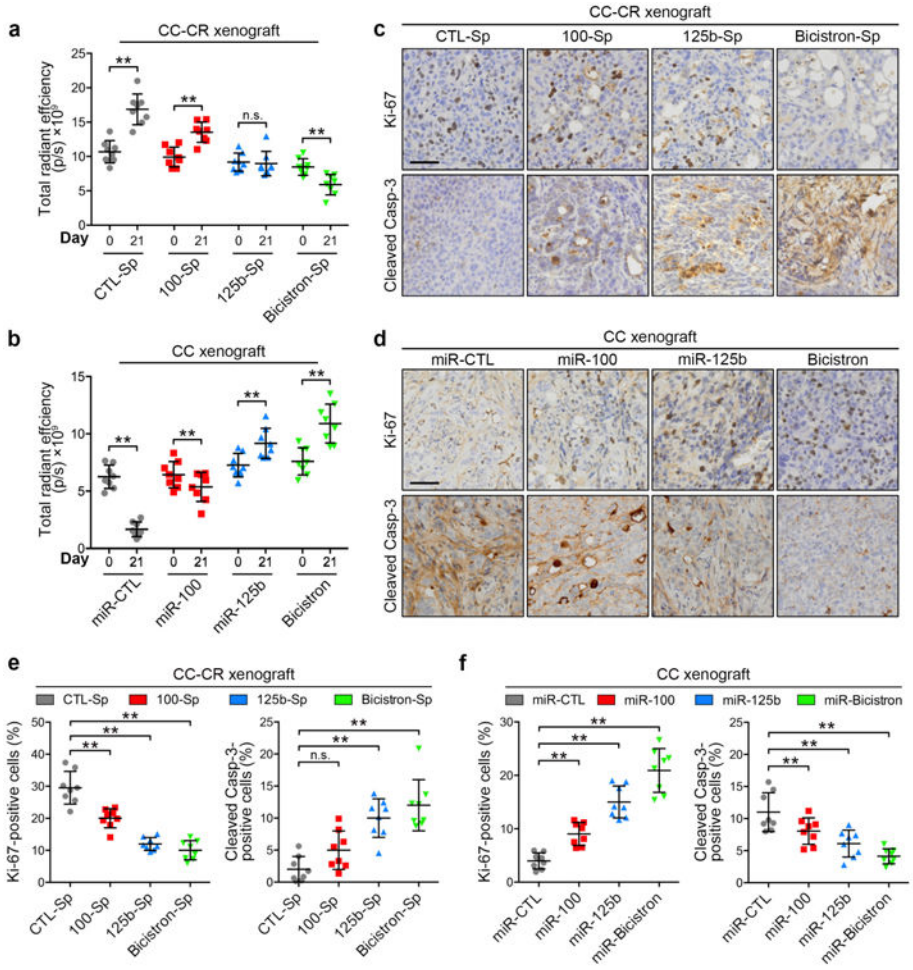


Extended Data Fig. 4.

miR-100 and miR-125b cooperativity drives cetuximab resistance in colorectal cancer and head and neck squamous cell cancer cell lines.

(a) Caco-2 cells stably overexpressing Bicistron or control (miR-CTL) were cultured in 3D for 5 days and treated with CTX (50 μg/ml) for 24 h. Immunofluorescence was performed for Cleaved Casp-3 (cyan) and Ki-67 (magenta) with quantification shown on the right. Scale bar, 50 μm. n=3 independent experiments. *P<0.05, **P<0.01 by Student's t test. **(b)** DLD-1 cells stably expressing Bicistron-Sp or control (CTL-Sp) were cultured in 3D for 10 days and treated with CTX (200 μg/ml) for 24 h. Staining of Cleaved Casp-3 (cyan) and Ki-67 (magenta) were shown. Scale bars, 50 μm. Quantification is shown on the right. n=3 independent experiments. *P<0.05, **P<0.01 by Student's t test. **(c, d)** Indicated cells were grown in 3D in normal medium (CTL) or treated with CTX (50 μg/ml for Caco-2, and 200 μg/ml for DLD-1) in 3D. The resultant colonies were counted. n=2 independent experiments performed in triplicate. **P<0.01 by one-way ANOVA followed by Dunnett's test compared with miR-CTL or CTL-Sp. **(e)** Left: CTX-R7 cells stably expressing miR-100 and/or miR-125b sponges were grown in normal medium (CTL) or treated with CTX (30 μg/ml). Cell viability was measured by cell counting kit-8 (CCK-8) assays after 72 h. n=3 independent experiments performed in triplicate. **P<0.01 by one-way ANOVA followed

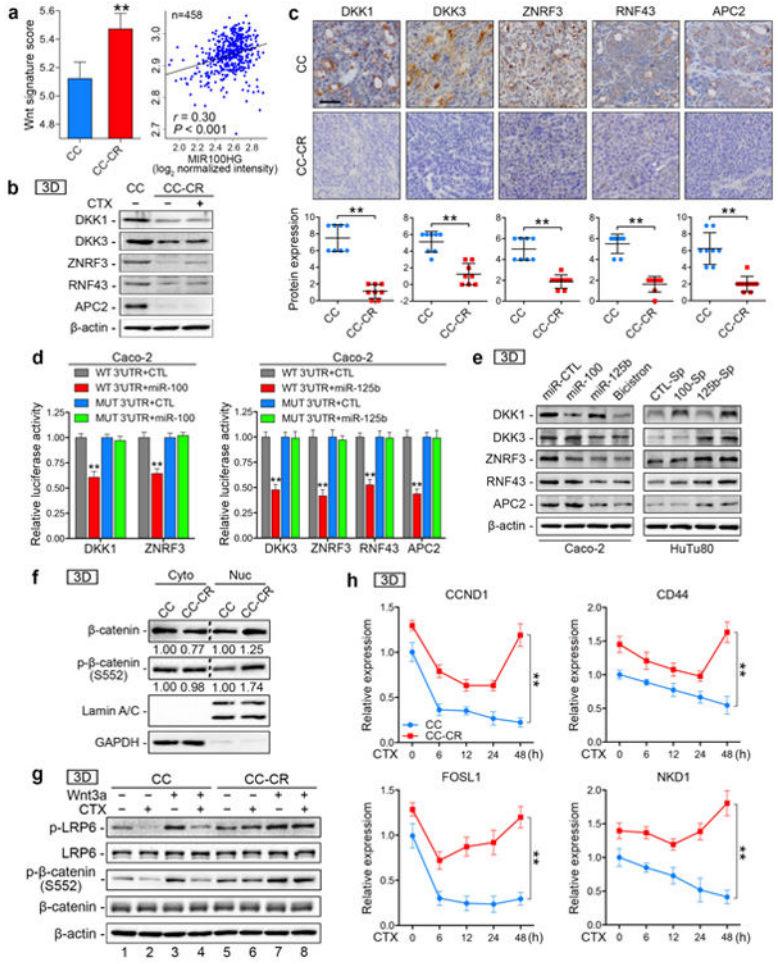
by Dunnett's test compared with CTL-Sp. Middle: qRT-PCR analysis of Wnt target genes in the stable bicistron sponge-transduced CTX-R7 cells. n=2 independent experiments performed in triplicate. *P<0.05, **P<0.01 by Student's t test. Right: CTX-R7 cells were treated with CTX (30 μg/ml) and/or ICG-001 (2 μM) for 72 h, and cell viability was measured by CCK-8 assays. n=2 independent experiments performed in triplicate. **P<0.01 by one-way ANOVA followed by by LSD post-hoc test. Data represent mean ± s.d. n.s., not significant.



Extended Data Fig. 5.

Effects of differential modulation of miR-100 and/or miR-125b on cetuximab responsiveness in CC and CC-CR *in vivo*.

(a, b) Quantification of radiant efficiency from tumors (n=8) represented on Fig. 3e and f. **P<0.01 by paired Student's t test. **(c-f)** Representative IHC images and quantification of Ki-67 and Cleaved Casp-3 from indicated xenografts (n=8) treated with CTX. Scale bars, 50 μm. **P<0.01 by one-way ANOVA followed by Dunnett's test in e and f. Data represent mean ± s.d. in a, b, e, and f. n.s., not significant.

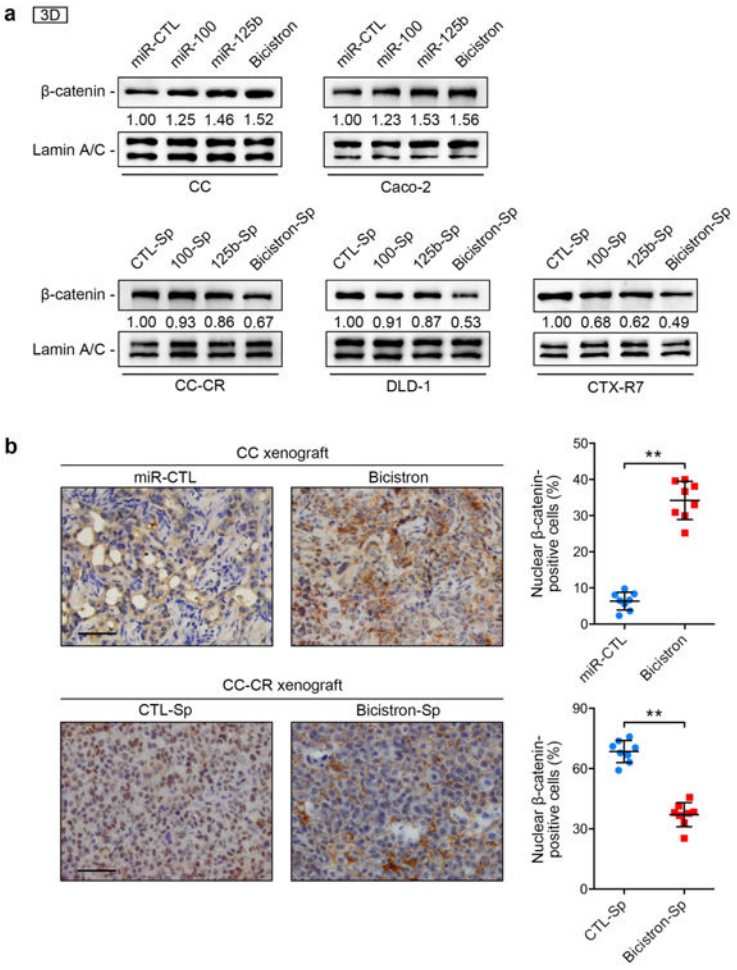


Extended Data Fig. 6.

miR-100/125b coordinately represses five Wnt/β-catenin negative regulators, resulting in increased Wnt signaling.

(a) Left: Wnt activation in CC and CC-CR cells was measured by the 64 Wnt/β -catenin target genes (Wnt signature score). ***P*<0.01 by Student's *t* test. Right: Scatter plots of MIR100HG expression versus 64-gene Wnt signature score on 458 CRC. Pearson correlation coefficients (*r*) and *P* values are shown. **(b)** Immunoblots of DKK1, DKK3, ZNRF3, RNF43, and APC2 levels from 3D cell lysates of CC and CC-CR. In CC-CR, cells were treated with CTX (CTX+, 3 μg/ml) or normal culture medium (CTX-) for 14 days in 3D before protein extraction. Representative of 3 independent experiments. **(c)** Top : representative IHC images of DKK1, DKK3, ZNRF3, RNF43, and APC2 in CC and CC-CR xenografts (n=8). Bottom: measurement of protein expression, ***P*<0.01 by Mann-Whitney *U* test. **(d)** Dual luciferase assays of genes predicted to be regulated by miR-100 or miR-125b in Caco-2. *Renilla* luciferase activity was normalized to firefly activity. n=2 independent experiments. ***P*<0.01 by Student's *t* test. **(e)** Immunoblots of indicated proteins in stable miRNA-transduced Caco-2 and sponge (Sp)-transduced HuTu80. Representative of 2 independent experiments. **(f)** Immunoblot of nuclear and cytoplasmic extracts for β-catenin and p-β-catenin (S552). Loading controls were GAPDH for

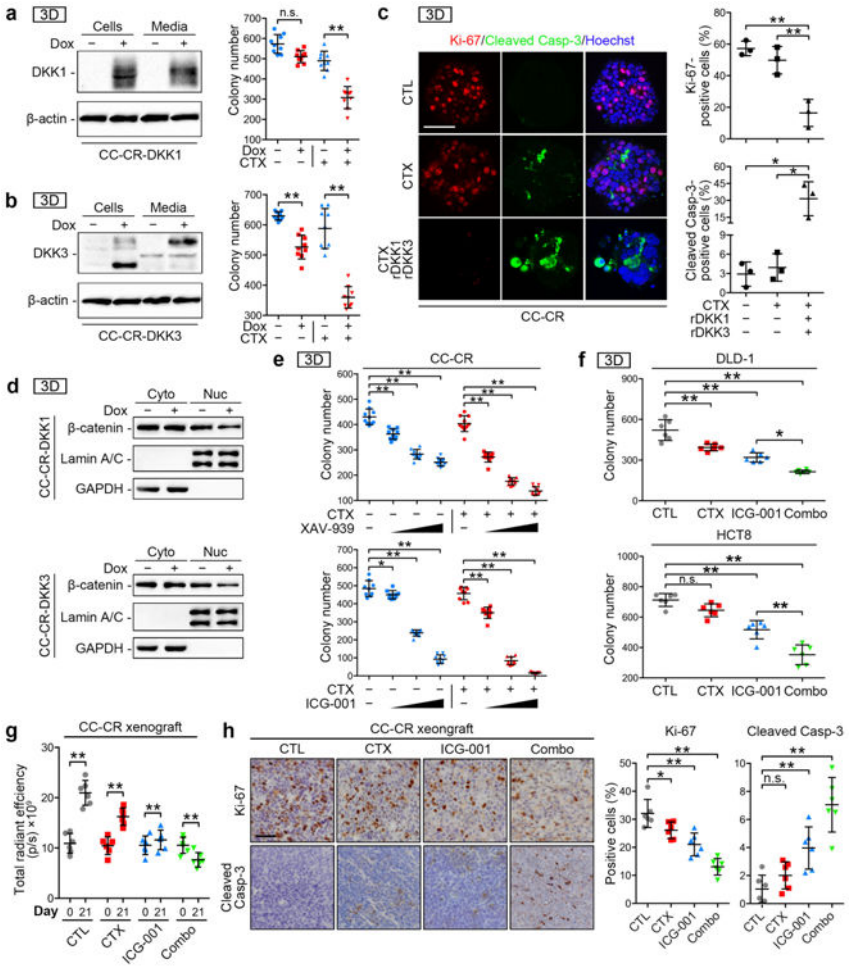
cytoplasmic fractions and Lamin A/C for nuclear fractions. **(g)** CC and CC-CR in 3D were treated with CTX (10 μg/ml) and/or Wnt3a (100 ng/ml). Immunoblots of indicated proteins after 48 h of treatment are shown. Representative of 3 independent experiments. **(h)** qRT-PCR analysis of Wnt targets CCND1, CD44, FOSL1, and NKD1 mRNAs at indicated time points following CTX (10 μg/ml) treatment in 3D. n=2 independent experiments performed in triplicate. ***P*<0.01 by two-way ANOVA test. Data represent mean ± s.d. in a, c, d, and h.



Extended Data Fig. 7.

Effects of differential modulation of miR-100 and/or miR-125b on nuclear β-catenin expression levels.

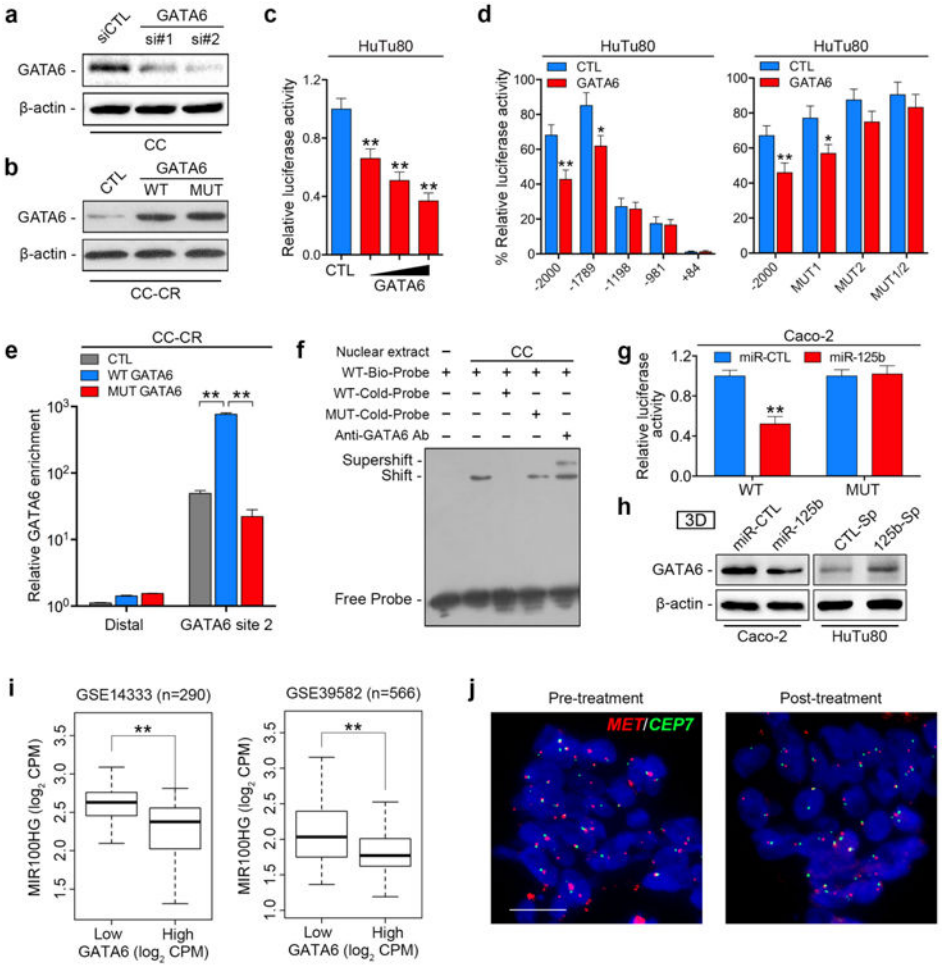
(a) Immunoblots for β-catenin from nuclear fractions in the CC and Caco-2 cells overexpressing miR-100 and/or miR-125b, or CC-CR, DLD-1 and CTX-R7 cells expressing miR-100 and/or miR-125b sponges. Lamin A/C served as the control for nuclear fractions. Representative of 2 independent experiments. **(b)** Representative IHC of β-catenin in the indicated xenografts (n=8). Scale bars: 50 μm. Quantification of nuclear β-catenin positive cells is shown. Data represent mean ± s.d. ***P*<0.01 by Student's *t* test.



Extended Data Fig. 8.

Blockade of Wnt signaling restores cetuximab responsiveness to cetuximab-resistant cells. **(a, b)** Left: CC-CR doxycycline (Dox)-on DKK1 or DKK3 cells were cultured in the presence or absence of Dox (1 μg/ml) and harvested at 48 h. Total cell lysates and conditioned media were harvested and subjected to immunoblot analysis. Right: indicated cells were grown in 3D in normal medium or treated with CTX (3 μg/ml). The resultant colonies were counted after 18 days. n=3 experiments performed in triplicate. $^{**}P<0.01$ by Student's *t* test. **(c)** CC-CR cells were grown in 3D in normal medium (CTL), treated with CTX (3 μg/ml) or in combination with recombinant DKK1 (rDKK1) and DKK3 (rDKK3) in 3D every 2 days. The resultant colonies were stained after 18 days for Cleaved Casp-3 (green) and Ki-67 (red). Scale bar, 50 μm. Quantification was shown. n=3 independent experiments. **(d)** Immunoblots for β-catenin from nuclear and cytoplasmic fractions of indicated cells upon CTX (10 μg/ml) treatment. Loading controls were GAPDH for cytoplasmic fractions and Lamin A/C for nuclear fractions. **(e)** CC-CR were treated with CTX (3 μg/ml), and/or XAV-939 (1, 5, 10 μM), and/or ICG-001 (1, 2.5, 5 μM) in 3D for 18 days, and colony number was determined. n=3 experiments performed in triplicate. **(f)** DLD-1 and HCT8 cells were treated with CTX (200 μg/ml) and/or ICG-001 (4 μM) for 14 days in 3D, and colony number was determined. n=2 independent experiments performed in

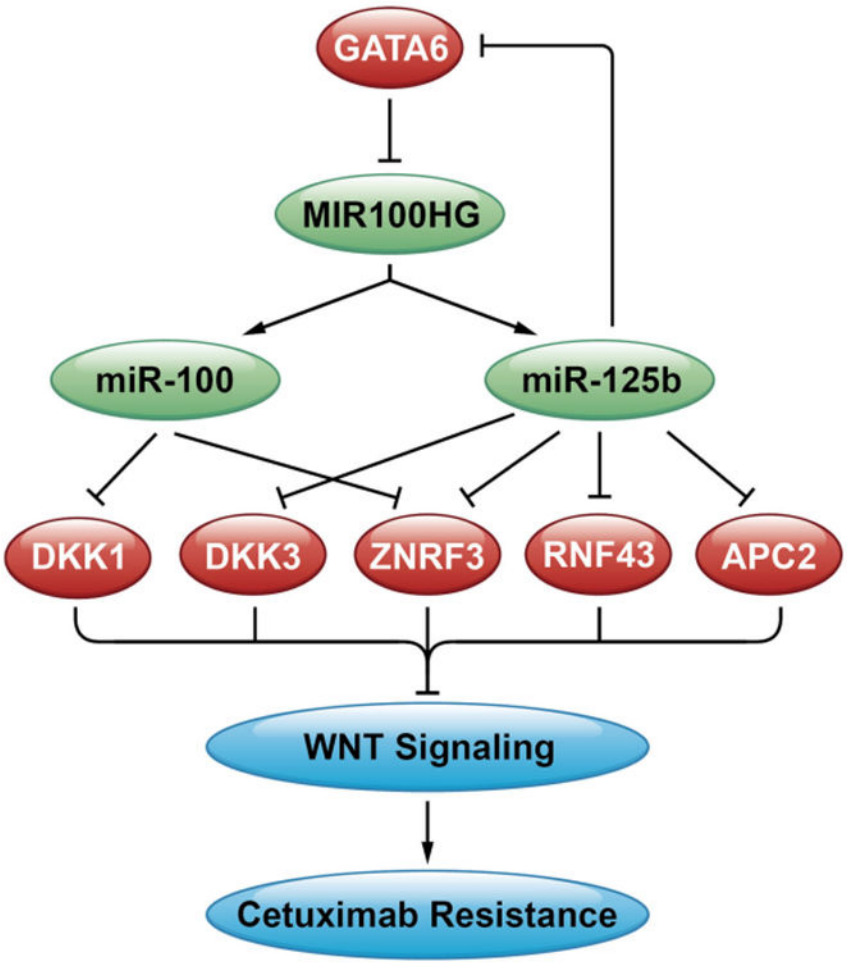
triplicate. **(g)** Quantification of radiant efficiency from tumors (n=6) represented on Fig. 4i. ***P*<0.01 by paired Student's *t* test. **(h)** Representative IHC images and quantification of Ki-67 and Cleaved Casp-3 from CC-CR xenografts (n=6) treated with control saline (CTL), or CTX (1 mg/mouse, i.p. injection, every 3 days), and/or ICG-001 (150 mg/kg, i.p. injection, daily). Scale bar, 50 μm. **P*<0.05, ***P*<0.01 by one-way ANOVA followed by Dunnett's test in c, e, and h, and one-way ANOVA followed by LSD post-hoc test in f. Data represent mean ± s.d. in a-c and e-h. n.s., not significant.



Extended Data Fig. 9.

A double-negative feedback loop between MIR100HG/miR-125b and GATA6. **(a)** Immunoblots of GATA6 expression in CC transfected with 2 independent siRNAs against GATA6 or control siRNA (siCTL). **(b)** Immunoblots of GATA6 expression in CC-CR transfected with either pcDNA3.1-GATA6 (WT GATA6), or pcDNA3.1-mutant GATA6 (MUT GATA6), or empty vector (CTL). **(c)** Luciferase reporter assays were performed in HuTu80 by co-transfection of pGL3-MIR100HG promoter luciferase reporter with increasing concentrations of pcDNA3.1-GATA6, and a *Renilla* control. Luciferase activity was normalized to *Renilla* values. n=3 independent experiments performed in triplicate. ***P*<0.01 by one-way ANOVA followed by Dunnett's test. **(d)** The luciferase vector pGL3

driven by either wild-type, deletion, or mutant (MUT) promoter was transfected in HuTu80, and luciferase activity was measured. n=3 independent experiments. *P<0.05, **P<0.01 by Student's t test. **(e)** ChIP assays were performed with anti-GATA6 antibody or control IgG in CC-CR overexpressing either WT GATA6, MUT GATA6, or CTL. The abundance of DNA within the MIR100HG promoter region was assessed by qRT-PCR with a primer pair spanning the GATA-binding site 2. A primer pair 6.4 kb distal to the MIR100HG promoter (Distal) was used as control. Data are presented as relative enrichment normalized to control IgG. **P<0.01 by one-way ANOVA followed by LSD post-hoc test. **(f)** EMSA using nuclear extracts from CC and the indicated probes. Ab, antibody. Representative of 3 independent experiments. **(g)** Luciferase reporter analysis of a wild-type (WT) or MUT GATA6 3′ UTR activity upon addition of synthetic miR-125b in Caco-2. n=2 independent experiments. **P<0.01 by Student's t test. **(h)** Immunoblots of GATA6 in stable miR-125b-transduced Caco-2 and 125b-Sp-transduced HuTu80. **(i)** Box plots showing MIR100HG expression in the lower (<25%) and the higher (>75%) quartiles of GATA6 expression from GEO CRC datasets GSE14333 and GSE39582. **P<0.01 by Mann–Whitney U test. **(j)** *MET* genomic status detected by FISH assay. There was no obvious change in *MET* copy number in 10 paired tumor specimens pre- and post-cetuximab treatment. Representative images are shown. Red, *MET* locus; green, chromosome 7 centromere (*CEP7*). Scale bar, 20 μm. Data represent mean ± s.d. in c-e and g.



Extended Data Fig. 10.

Model of a new mode of acquired and *de novo* cetuximab resistance. We propose a complex circuitry in which the lncRNA MIR100HG through embedded miR-100 and miR-125b confers cetuximab resistance by targeting and decreasing expression of five negative regulators of Wnt signaling, DKK1, DKK3, ZNRF3, RNF43, and APC2. This results in increased Wnt signaling and cetuximab resistance; this resistance can be overcome by blockade of Wnt signaling. We present evidence that GATA6 represses MIR100HG expression, but that miR-125b targets GATA6 to relieve this repression.

Figure 1e

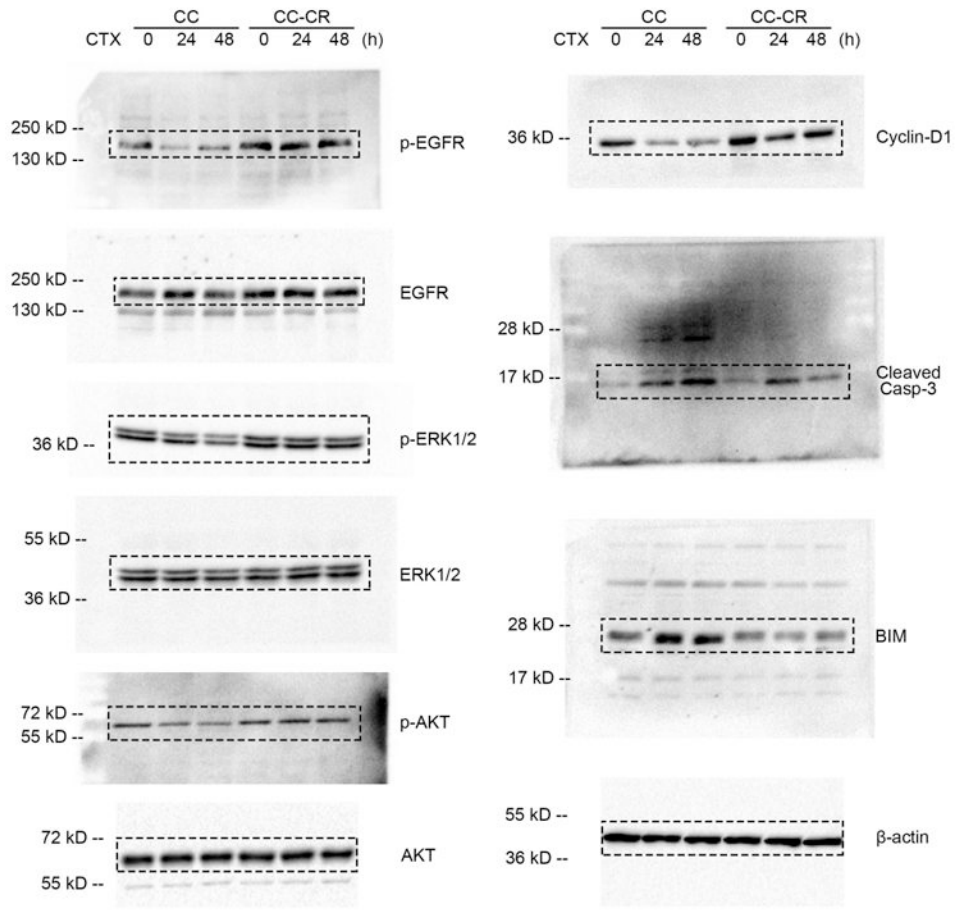


Figure 4d

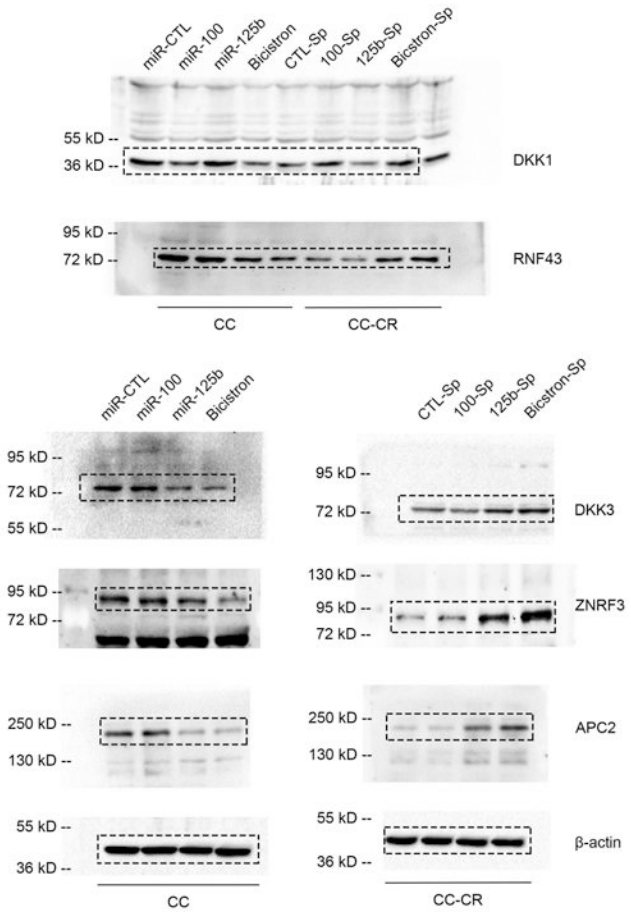


Figure 5a

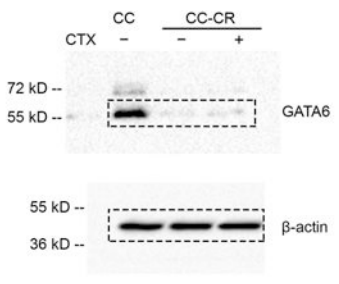
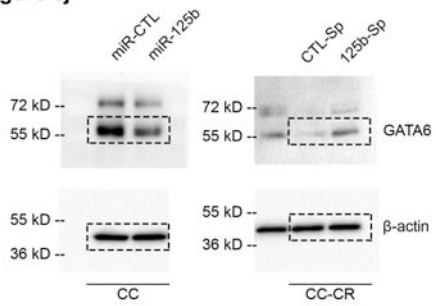
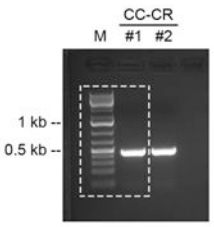


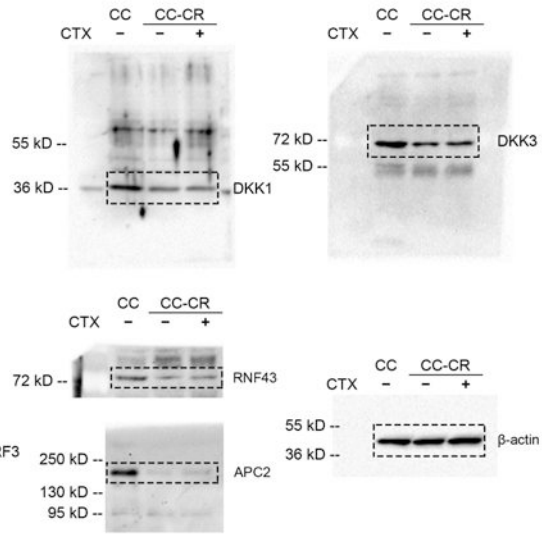
Figure 5j

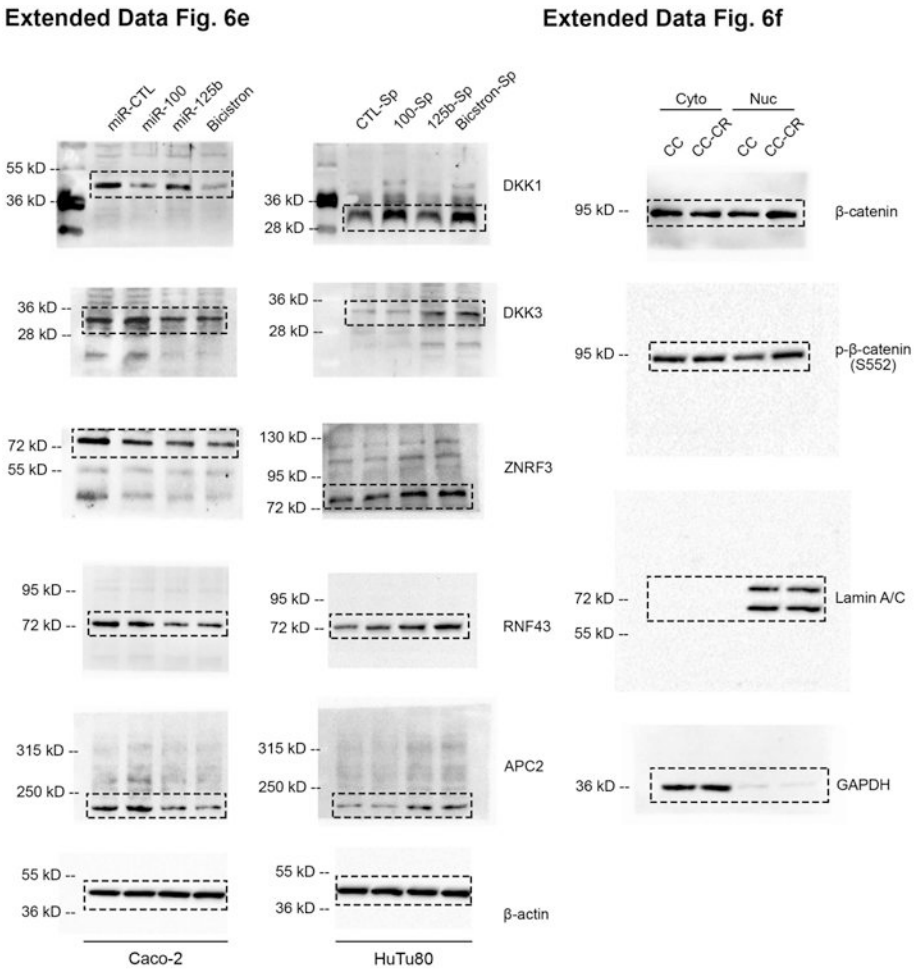


Extended Data Fig. 2c

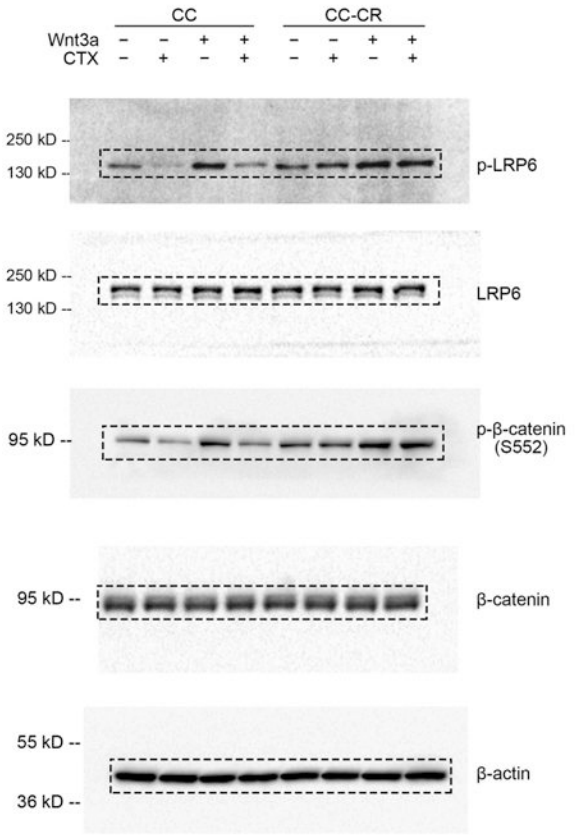


Extended Data Fig. 6b

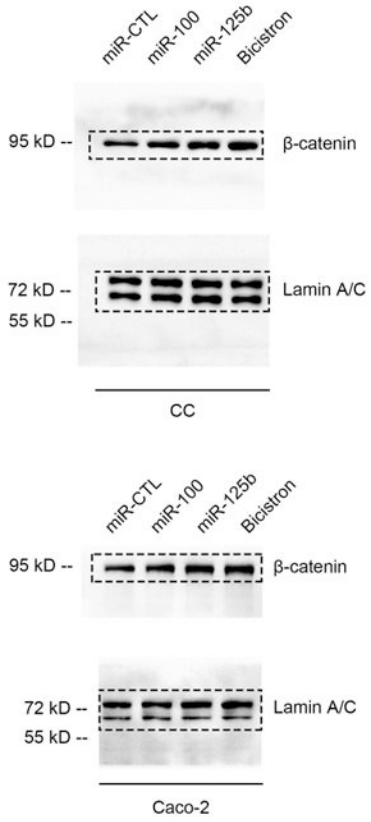




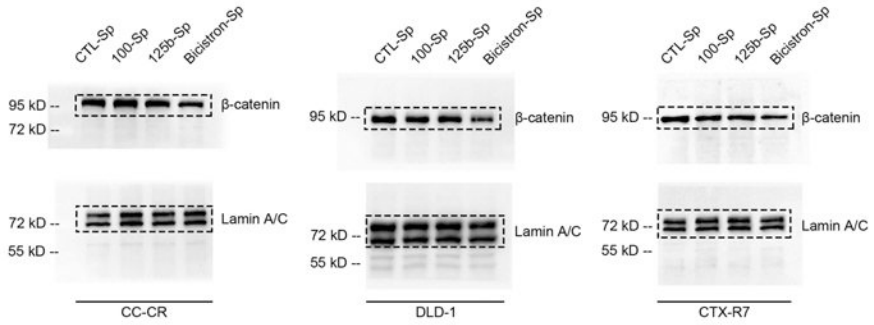
Extended Data Fig. 6g



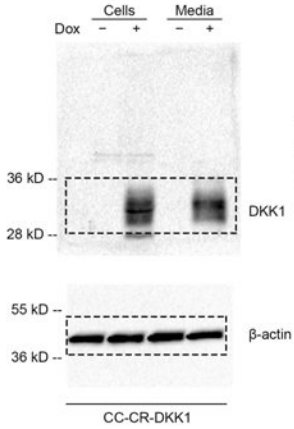
Extended Data Fig. 7a



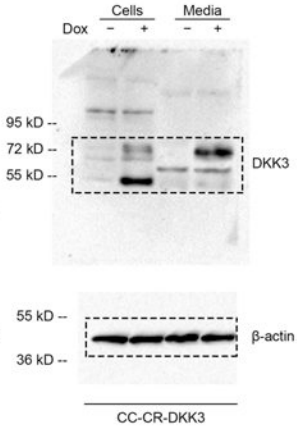
Extended Data Fig. 7a



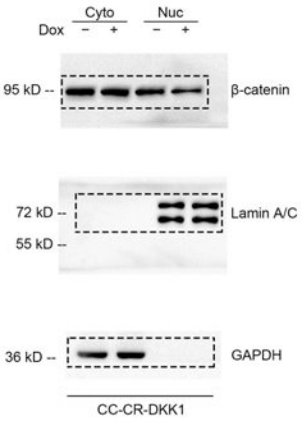
Extended Data Fig. 8a

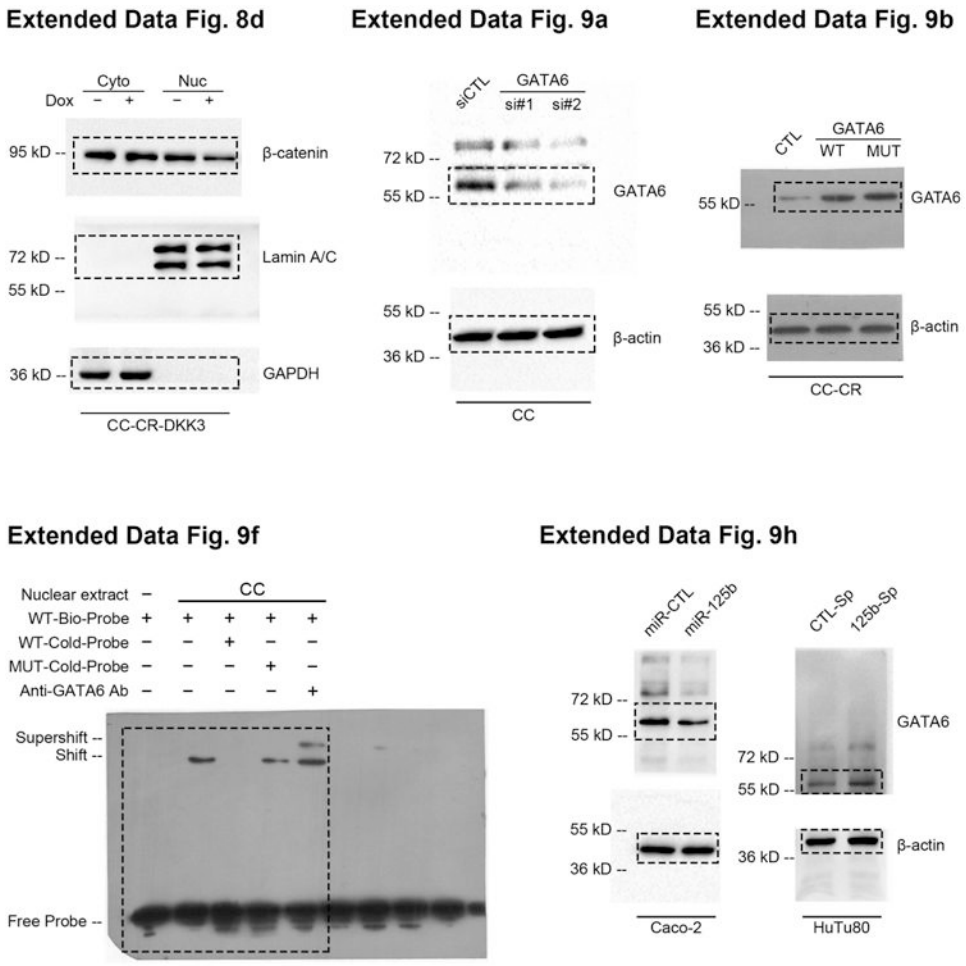


Extended Data Fig. 8b



Extended Data Fig. 8d





Extended Data Fig. 11.
Original images of immunoblots with molecular weight standards.

Extended Data Table 1

Chr	Pos	Ref	Alt	Type
chr1	1.17E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chr16	74537530	C	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chrX	48762192	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chrX	1.34E+08	C	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chrX	1.35E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr1	1254773	G	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chr1	17740089	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chr1	32653705	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr1	47024318	G	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODING
chr1	51210391	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr1	93303020	C	T	STOP_GAINED(HIGH)
chr1	1.46E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE)

Chr	Pos	Ref	Alt	Type
chr1	1.55E+08	T	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr1	1.86E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr1	2.27E+08	C	A	NON_SYNONYMOUS_CODING(MODERATE)
chr1	2.49E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr2	56420411	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr2	65541153	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr2	84668385	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr2	98999879	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr2	99797317	T	A	NON_SYNONYMOUS_CODING(MODERATE)
chr2	1.44E+08	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr2	1.6E+08	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr2	1.98E+08	A	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr2	2.36E+08	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr3	30713286	G	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr3	48604415	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr3	48604422	C	A	NON_SYNONYMOUS_CODING(MODERATE)
chr3	49689732	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr3	49765037	G	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr3	75788432	T	A	NON_SYNONYMOUS_CODING(MODERATE)
chr3	1.01E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr3	1.88E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr3	1.9E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr4	5577982	T	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr4	74021398	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr4	89709004	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr4	1.26E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr4	1.55E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr4	1.56E+08	T	A	NON_SYNONYMOUS_CODING(MODERATE)
chr4	1.56E+08	C	T	SPLICE_SITE_DONOR(HIGH)
chr5	434878	G	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr5	23527585	GACACACACA	GACACACACACA	FRAME_SHIFT(HIGH)
chr5	31799521	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr5	54624530	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr5	1.32E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr5	1.4E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr5	1.4E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr5	1.41E+08	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr5	1.79E+08	C	G	NON_SYNONYMOUS_CODING(MODERATE)
chr6	553900	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr6	25972305	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr6	39073484	G	T	STOP_GAINED(HIGH)
chr6	43267448	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC

Chr	Pos	Ref	Alt	Type
chr6	1.61E+08	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr6	1.71E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr7	1538645	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr7	48349667	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr7	91603099	A	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr7	91756945	C	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr7	97863036	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr7	1E+08	T	TG	FRAME_SHIFT(HIGH);FRAME_SHIFT(HIGH)
chr7	1.05E+08	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr7	1.27E+08	C	A	STOP_GAINED(HIGH)
chr7	1.42E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr7	1.49E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr8	623772	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr8	35453086	T	G	NON_SYNONYMOUS_CODING(MODERATE)
chr8	1.42E+08	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr8	1.45E+08	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr9	34255869	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr9	86452279	A	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr9	96052320	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.16E+08	T	G	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.25E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.28E+08	C	G	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.34E+08	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.34E+08	T	A	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.37E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr9	1.39E+08	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr10	17199504	G	A	STOP_GAINED(HIGH)
chr10	26455049	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr10	32304529	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr10	50819232	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr10	98129916	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr10	1.23E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr10	1.35E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr11	612778	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr11	773623	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr11	1078307	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr11	2427314	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr11	66619985	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr11	77378107	G	C	NON_SYNONYMOUS_CODING(MODERATE)
chr11	1.19E+08	C	A	NON_SYNONYMOUS_CODING(MODERATE)
chr11	1.19E+08	A	T	STOP_GAINED(HIGH);STOP_GAINED(HIGH)
chr12	6933276	C	A	NON_SYNONYMOUS_CODING(MODERATE)

Chr	Pos	Ref	Alt	Type
chr12	9260161	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr12	53453359	G	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr12	58197052	A	T	NON_SYNONYMOUS_CODING(MODERATE)
chr12	69126435	G	T	NON_SYNONYMOUS_CODING(MODERATE)
chr12	77235849	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr13	39425170	C	T	STOP_GAINED(HIGH)
chr14	73957946	T	G	NON_SYNONYMOUS_CODING(MODERATE)
chr14	1.04E+08	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr15	66601046	G	T	SPLICE_SITE_ACCEPTOR(HIGH);SPLICE_SITE_ACCEPTOR(HIGH)
chr15	90801386	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr16	613451	T	C	NON_SYNONYMOUS_CODING(MODERATE)
chr16	31230697	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr17	4167225	A	G	SPLICE_SITE_DONOR(HIGH);SPLICE_SITE_DONOR(HIGH)
chr17	8045692	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr17	37565110	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr17	62856060	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr17	73499968	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr17	76449531	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr17	79555984	C	A	NON_SYNONYMOUS_CODING(MODERATE)
chr18	3272991	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr18	29867458	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr19	1295566	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr19	9059365	C	A	STOP_GAINED(HIGH)
chr19	16982170	CTT	CTTT	FRAME_SHIFT(HIGH)
chr19	19414563	A	G	NON_SYNONYMOUS_CODING(MODERATE)
chr19	48305970	G	A	NON_SYNONYMOUS_CODING(MODERATE)
chr19	48922916	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr19	49377522	T	G	STOP_GAINED(HIGH)
chr19	54759960	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr19	56185411	T	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr19	57325702	C	T	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr19	57640078	T	A	NON_SYNONYMOUS_CODING(MODERATE)
chr20	17928162	T	A	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr20	31656753	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr20	43052973	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr20	48808388	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr20	50342443	A	C	NON_SYNONYMOUS_CODING(MODERATE)
chr20	54970702	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr21	41559073	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr21	43867183	T	C	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr21	45994548	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr21	46086757	G	A	NON_SYNONYMOUS_CODING(MODERATE)

Chr	Pos	Ref	Alt	Type
chr21	46086758	C	T	NON_SYNONYMOUS_CODING(MODERATE)
chr22	23596000	A	G	NON_SYNONYMOUS_CODING(MODERATE);NON_SYNONYMOUS_CODINC
chr3	15613277	CTTAA	C	SPLICE_SITE_ACCEPTOR(HIGH)

Extended Data Table 2

Relative expression of ERBB1-4, seven EGFR ligands, and MET in CC and CC-CR

Gene Name	FPKM[1] CC 1	FPKM CC 2	FPKM CC 3	FPKM CC-CR 1	FPKM CC-CR 2	FPKM CC-CR 3	Log_2FC^2 (CC-CR vs CC)	FDR
EGFR	7.113792	7.216315	6.9096	10.53567	8.838911	9.740446	0.457468	6.32E-0
ERBB2	24.23383	24.52192	23.97762	23.07252	25.54938	21.22859	-0.05656	0.73088
ERBB3	18.79607	18.91252	17.90869	19.07041	20.00568	14.89755	-0.04198	0.85980
ERBB4	0	0.005107	0	0.016069	0	0.004937	NA	NA
AREG	21.5716	21.76646	20.76116	13.48585	13.73754	17.10571	-0.5276	0.00061
BTC	6.413408	5.307637	6.723783	5.5914	5.375864	8.232057	0.062581	0.88137
EGF	0.568842	0.769771	0.574685	0.488951	0.322759	0.366821	-0.69627	0.01236
EPGN	0.139733	0.109859	0.183119	0	0.263141	0.176997	NA	NA
EREG	31.82304	31.86035	37.40712	21.31407	23.27111	33.68543	-0.36529	0.11793
HBEGF	3.170548	2.707305	2.575539	3.416546	3.66036	3.297967	0.295692	0.17501
TGFA	13.67483	12.79573	11.80663	19.20377	16.29553	15.21587	0.407307	0.00045
MET	135.2686	157.1884	149.2257	213.2824	228.8418	221.7658	0.590334	3.12E-0

[1]FPKM stands for Fragments Per Kilobase of transcript per Million mapped reads.

[2]Differential expression analysis was performed on RNA-Seq counts using edgeR.

Extended Data Table 3

Mutational status of 30 CRC cell lines used in Fig. 2f and their response to cetuximab

Cell line	Mutational status	Mean CTX inhibition rate (%)	CTX response category
NCI-H508	*BRAF*	83.4[1]	sensitive
V9P	WT	82.0[3]	sensitive
DiFi	WT	80.8[1]	sensitive
LIM1215	WT	79.6[2]	sensitive
GEO	*KRAS*	68.7[2]	sensitive
SW403	*KRAS*	66.0[2]	sensitive
SNUC4	WT	48.3[1]	partially responsive
Caco-2	WT	47.7[2]	partially responsive
SW948	*KRAS*	42.7[2]	partially responsive
HT29	*BRAF*	36.5[2]	partially responsive
SK-CO-1	*KRAS*	33.9[2]	partially responsive
DLD-1	*KRAS*	24.9[2]	resistant
SW480	*KRAS*	23.7[2]	resistant

Cell line	Mutational status	Mean CTX inhibition rate (%)	CTX response category
SW837	*KRAS*	21.8[2]	resistant
SW48	WT	21.8[2]	resistant
SW620	*KRAS*	14.5[2]	resistant
LoVo	*KRAS*	14.3[1]	resistant
COLO205	*BRAF*	12.5[1]	resistant
T84	*KRAS*	11.2[2]	resistant
LS174T	*KRAS*	9.7[1]	resistant
NCI-H716	WT	9.7[1]	resistant
HCT8	*KRAS*	8.4[1]	resistant
HCT15	*KRAS*	4.5[2]	resistant
SW1116	*KRAS*	2.0[1]	resistant
LIM2405	*KRAS*	2.0[2]	resistant
RKO	*BRAF*	0.4[1]	resistant
COLO320DM	WT	-3.2[1]	resistant
HuTu80	WT	-5.4[1]	resistant
LS123	*KRAS*	-4.8[1]	resistant
HCT116	*KRAS*	-14.1[2]	resistant

[1]Data came from Medico, E. *et al*, *Nat Commun*, 2015[17].

[2]Data came from Jhawer, M. *et al*, *Cancer Res*, 2008[16].

[3]Experimental data from the present study.

Extended Data Table 4

miRNA	PathName	PathFg	PathBg	GenomeFG	GenomeBG	pval	BH
hsa-miR-100-5p	Pathways in cancer	197	330	8828	19747	2.49E-08	4.84E-06
hsa-miR-100-5p	Pancreatic cancer	52	75	8828	19747	1.39E-05	0.00264553
hsa-miR-100-5p	Wnt signaling pathway	94	152	8828	19747	1.49E-05	0.00285334
hsa-miR-100-5p	mTOR signaling pathway	39	53	8828	19747	1.87E-05	0.00354634
hsa-miR-100-5p	Ubiquitin mediated proteolysis	83	134	8828	19747	4.24E-05	0.00797254
hsa-miR-100-5p	Glioma	45	65	8828	19747	5.51E-05	0.01025001
hsa-miR-100-5p	Endocytosis	110	187	8828	19747	6.79E-05	0.0124941
hsa-miR-100-5p	Chronic myeloid leukemia	50	75	8828	19747	0.00010082	0.01834838
hsa-miR-100-5p	Calcium signaling pathway	104	178	8828	19747	0.00015222	0.02739988
hsa-miR-100-5p	Focal adhesion	116	203	8828	19747	0.00023268	0.04141736
hsa-miR-100-5p	Leukocyte transendothelial migration	71	116	8828	19747	0.00024812	0.04416472
hsa-miR-100-5p	Insulin signaling pathway	83	139	8828	19747	0.00025507	0.04514807
hsa-miR-100-5p	Non small cell lung cancer	37	54	8828	19747	0.00034689	0.06035915
hsa-miR-100-5p	Apoptosis	55	87	8828	19747	0.00038021	0.06577642
hsa-miR-100-5p	MAPK signaling pathway	149	272	8828	19747	0.0004966	0.0849182
hsa-miR-100-5p	Renal cell carcinoma	46	71	8828	19747	0.00050651	0.08661247
hsa-miR-100-5p	Long term potentiation	46	71	8828	19747	0.00050651	0.08661247
hsa-miR-100-5p	Axon guidance	76	129	8828	19747	0.00079389	0.13257995
hsa-miR-100-5p	Small cell lung cancer	52	84	8828	19747	0.00110523	0.17904734

miRNA	PathName	PathFg	PathBg	GenomeFG	GenomeBG	pval	BH
hsa-miR-100-5p	Colorectal cancer	53	86	8828	19747	0.00115401	0.18694947
hsa-miR-100-5p	Aldosterone regulated sodium reabsorption	29	42	8828	19747	0.00124064	0.19974264
hsa-miR-100-5p	Phosphatidylinositol signaling system	47	76	8828	19747	0.00193595	0.30200877
hsa-miR-100-5p	Chondroitin sulfate biosynthesis	17	22	8828	19747	0.00195406	0.30483395
hsa-miR-100-5p	VEGF signaling pathway	48	78	8828	19747	0.00201937	0.31502171
hsa-miR-100-5p	Hedgehog signaling pathway	36	56	8828	19747	0.0024547	0.37556932
hsa-miR-100-5p	Cell adhesion molecules CAMs	76	133	8828	19747	0.0025693	0.39310273
hsa-miR-100-5p	Vascular smooth muscle contraction	67	116	8828	19747	0.0031307	0.47273626
hsa-miR-100-5p	Glycosphingolipid biosynthesis lacto and neolacto series	19	26	8828	19747	0.00320356	0.48373791
hsa-miR-100-5p	Amyotrophic lateral sclerosis ALS	35	55	8828	19747	0.00359894	0.53624172
hsa-miR-100-5p	Fc gamma R mediated phagocytosis	57	97	8828	19747	0.00366845	0.54659853
hsa-miR-100-5p	Lysosome	69	121	8828	19747	0.0042201	0.61613498
hsa-miR-100-5p	Dilated cardiomyopathy	55	94	8828	19747	0.00484392	0.6975252
hsa-miR-100-5p	Melanoma	43	71	8828	19747	0.00514181	0.73527865
hsa-miR-100-5p	Heparan sulfate biosynthesis	18	26	8828	19747	0.01012535	1
hsa-miR-100-5p	Lysine degradation	28	45	8828	19747	0.01350559	1
hsa-miR-100-5p	Melanogenesis	57	102	8828	19747	0.01500276	1
hsa-miR-100-5p	GnRH signaling pathway	56	105	8828	19747	0.04641394	1
hsa-miR-100-5p	Arrhythmogenic right ventricular cardiomyopathy ARVC	41	74	8828	19747	0.04152065	1
hsa-miR-100-5p	Keratan sulfate biosynthesis	11	15	8828	19747	0.02407578	1
hsa-miR-100-5p	Regulation of actin cytoskeleton	110	212	8828	19747	0.02068052	1
hsa-miR-100-5p	Type II diabetes mellitus	29	49	8828	19747	0.02922526	1
hsa-miR-100-5p	Neurotrophin signaling pathway	70	129	8828	19747	0.0180486	1
hsa-miR-100-5p	O Glycan biosynthesis	20	30	8828	19747	0.01266537	1
hsa-miR-100-5p	Valine leucine and isoleucine degradation	28	45	8828	19747	0.01350559	1
hsa-miR-100-5p	Adherens junction	45	76	8828	19747	0.0076372	1
hsa-miR-100-5p	ABC transporters	26	44	8828	19747	0.03875465	1
hsa-miR-100-5p	Basal cell carcinoma	34	55	8828	19747	0.00786933	1
hsa-miR-100-5p	Prostate cancer	48	89	8828	19747	0.05007106	1
hsa-miR-100-5p	TGF beta signaling pathway	47	86	8828	19747	0.04039675	1
hsa-miR-100-5p	Dorso ventral axis formation	16	24	8828	19747	0.02509059	1
hsa-miR-100-5p	Nicotinate and nicotinamide metabolism	17	24	8828	19747	0.0087657	1
hsa-miR-100-5p	Adipocytokine signaling pathway	41	70	8828	19747	0.01352383	1
hsa-miR-100-5p	Fc epsilon RI signaling pathway	46	82	8828	19747	0.0248563	1
hsa-miR-100-5p	Butanoate metabolism	21	35	8828	19747	0.04970063	1
hsa-miR-100-5p	Notch signaling pathway	28	47	8828	19747	0.02864452	1
hsa-miR-100-5p	T cell receptor signaling pathway	58	110	8828	19747	0.05508698	1
hsa-miR-100-5p	p53 signaling pathway	40	68	8828	19747	0.01330284	1
hsa-miR-100-5p	ErbB signaling pathway	51	89	8828	19747	0.0112398	1

miRNA	PathName	PathFg	PathBg	GenomeFG	GenomeBG	pval	BH
hsa-miR-100-5p	Hypertrophic cardiomyopathy HCM	49	86	8828	19747	0.01467435	1
miRNA	PathName	PathFg	PathBg	GenomeFG	GenomeBG	pval	BH
hsa-miR-125b-5p	MAPK signaling pathway	217	272	12145	19747	5.52E-11	1.07E-08
hsa-miR-125b-5p	Axon guidance	108	129	12145	19747	3.01E-08	5.77E-06
hsa-miR-125b-5p	Pathways in cancer	249	330	12145	19747	4.35E-08	8.35E-06
hsa-miR-125b-5p	Regulation of actin cytoskeleton	165	212	12145	19747	2.56E-07	4.84E-05
hsa-miR-125b-5p	Insulin signaling pathway	113	139	12145	19747	3.55E-07	6.71E-05
hsa-miR-125b-5p	Glioma	58	65	12145	19747	5.88E-07	0.0001106
hsa-miR-125b-5p	Cell adhesion molecules CAMs	108	133	12145	19747	7.17E-07	0.00013411
hsa-miR-125b-5p	Wnt signaling pathway	121	152	12145	19474	1.1E-06	0.00022601
hsa-miR-125b-5p	ErbB signaling pathway	74	89	12145	19747	7.66E-06	0.0014179
hsa-miR-125b-5p	Hedgehog signaling pathway	49	56	12145	19747	1.64E-05	0.00298105
hsa-miR-125b-5p	Endocytosis	142	187	12145	19747	1.87E-05	0.00339254
hsa-miR-125b-5p	Neurotrophin signaling pathway	101	129	12145	19747	3.36E-05	0.00602008
hsa-miR-125b-5p	Chronic myeloid leukemia	62	75	12145	19747	6.00E-05	0.01061902
hsa-miR-125b-5p	Pancreatic cancer	62	75	12145	19747	6.00E-05	0.01061902
hsa-miR-125b-5p	Melanoma	59	71	12145	19747	6.69E-05	0.01177116
hsa-miR-125b-5p	Focal adhesion	151	203	12145	19747	6.80E-05	0.01196783
hsa-miR-125b-5p	Notch signaling pathway	41	47	12145	19747	9.72E-05	0.0170169
hsa-miR-125b-5p	Non small cell lung cancer	46	54	12145	19747	0.0001296	0.02255083
hsa-miR-125b-5p	Leukocyte transendothelial migration	90	116	12145	19747	0.00016225	0.02774454
hsa-miR-125b-5p	Long term potentiation	58	71	12145	19747	0.00019926	0.03367434
hsa-miR-125b-5p	Colorectal cancer	68	86	12145	19747	0.00037718	0.06185754
hsa-miR-125b-5p	Calcium signaling pathway	131	178	12145	19747	0.00043597	0.07106345
hsa-miR-125b-5p	Glycosphingolipid biosynthesis lacto and neolacto series	24	26	12145	19747	0.00046625	0.07596159
hsa-miR-125b-5p	mTOR signaling pathway	44	53	12145	19747	0.00059905	0.09644727
hsa-miR-125b-5p	Endometrial cancer	43	52	12145	19747	0.00081681	0.12905666
hsa-miR-125b-5p	Phosphatidylinositol signaling system	60	76	12145	19747	0.00088761	0.14024252
hsa-miR-125b-5p	Prostate cancer	69	89	12145	19747	0.00095427	0.14982021
hsa-miR-125b-5p	VEGF signaling pathway	61	78	12145	19747	0.00125492	0.19576823
hsa-miR-125b-5p	T cell receptor signaling pathway	83	110	12145	19747	0.0013653	0.21298668
hsa-miR-125b-5p	Vascular smooth muscle contraction	87	116	12145	19747	0.00146085	0.22643156
hsa-miR-125b-5p	Lysosome	90	121	12145	19747	0.00189772	0.28891677
hsa-miR-125b-5p	Long term depression	57	73	12145	19747	0.00192366	0.29239561
hsa-miR-125b-5p	Melanogenesis	77	102	12145	19747	0.00196211	0.29824123
hsa-miR-125b-5p	GnRH signaling pathway	79	105	12145	19747	0.00203279	0.30695151
hsa-miR-125b-5p	Adherens junction	59	76	12145	19747	0.00207114	0.31274149
hsa-miR-125b-5p	Basal cell carcinoma	44	55	12145	19747	0.00265051	0.38962493
hsa-miR-125b-5p	Renal cell carcinoma	55	71	12145	19747	0.00316848	0.45635002
hsa-miR-125b-5p	Apoptosis	66	87	12145	19747	0.00326215	0.46975023
hsa-miR-125b-5p	Chemokine signaling pathway	134	189	12145	19747	0.00422669	0.59596343

miRNA	PathName	PathFg	PathBg	GenomeFG	GenomeBG	pval	BH
hsa-miR-125b-5p	Fc epsilon RI signaling pathway	62	82	12145	19747	0.00490539	0.68675477
hsa-miR-125b-5p	Type II diabetes mellitus	39	49	12145	19747	0.00543661	0.75568815
hsa-miR-125b-5p	Fc gamma R mediated phagocytosis	72	97	12145	19747	0.00565361	0.78585178
hsa-miR-125b-5p	Arrhythmogenic right ventricular cardiomyopathy ARVC	56	74	12145	19747	0.00711268	0.98155051
hsa-miR-125b-5p	Epithelial cell signaling in Helicobacter pylori infection	52	71	12145	19747	0.0257727	1
hsa-miR-125b-5p	Chondroitin sulfate biosynthesis	18	22	12145	19747	0.03634967	1
hsa-miR-125b-5p	SNARE interactions in vesicular transport	30	39	12145	19747	0.03157198	1
hsa-miR-125b-5p	Glycerophospholipid metabolism	51	70	12145	19747	0.03130765	1
hsa-miR-125b-5p	O Glycan biosynthesis	25	30	12145	19747	0.00882287	1
hsa-miR-125b-5p	TGF beta signaling pathway	62	86	12145	19747	0.02614341	1
hsa-miR-125b-5p	p53 signaling pathway	50	68	12145	19747	0.02553146	1
hsa-miR-125b-5p	Hypertrophic cardiomyopathy HCM	62	86	12145	19747	0.02614341	1
hsa-miR-125b-5p	Gap junction	64	90	12145	19747	0.03662525	1
hsa-miR-125b-5p	Ubiquitin mediated proteolysis	95	134	12145	19747	0.01448207	1
hsa-miR-125b-5p	Inositol phosphate metabolism	41	54	12145	19747	0.01834602	1
hsa-miR-125b-5p	Keratan sulfate biosynthesis	13	15	12145	19747	0.03506366	1
hsa-miR-125b-5p	Acute myeloid leukemia	43	58	12145	19747	0.03011936	1
hsa-miR-125b-5p	B cell receptor signaling pathway	54	75	12145	19747	0.0377669	1
hsa-miR-125b-5p	Tight junction	94	132	12145	19747	0.01241454	1
hsa-miR-125b-5p	Bladder cancer	33	43	12145	19747	0.02592105	1
hsa-miR-125b-5p	Vibrio cholerae infection	40	55	12145	19747	0.05536572	1
hsa-miR-125b-5p	Small cell lung cancer	61	84	12145	19747	0.02174543	1
hsa-miR-125b-5p	Adipocytokine signaling pathway	52	70	12145	19747	0.0169451	1
hsa-miR-125b-5p	Dilated cardiomyopathy	66	94	12145	19747	0.04946503	1

Extended Data Table 5

Schema of miR-100 and/or miR-125b binding sites in predicted target 3′ UTR sequences of human genes

Gene	Target Site[1]	3′ UTR Position
DKK1		88-98
DKK1		422-421
DKK3		641-647
ZNRF3		1164-1172
ZNRF3		475-481

Gene	Target Site[1]	3′ UTR Position
ZNRF3		1529-1535
RNF43		1308-1314
RNF43		1462-1468
APC2		625-631
APC2		1612-1618
GATA6		851-857

[1]Putative binding sites of miR-100 or miR-125b were mutated and highlighted in red.

Extended Data Table 6
Predicted transcription factors for MIR100HG

Feature	Name[1]	logFC[2]	logCPM	LR	*p*-Value	FDR
ENSG00000141448	GATA6	-5.4301	3.5893	538.9645	0.0000	0.0000
ENSG00000156127	BATF	-4.9970	-0.5743	32.7572	0.0000	0.0000
ENSG00000089225	TBX5	-4.7068	-0.6593	31.4959	0.0000	0.0000
ENSG00000054598	FOXC1	-1.2082	3.0427	42.6754	0.0000	0.0000
ENSG00000007372	PAX6	-0.8470	2.0843	21.5947	0.0000	0.0001
ENSG00000160973	FOXH1	-0.7994	0.8274	6.3042	0.0120	0.0741
ENSG00000114315	HES1	-0.7992	1.4362	10.5755	0.0011	0.0124
ENSG00000150907	FOXO1	0.6164	4.8378	32.5211	0.0000	0.0000
ENSG00000185630	PBX1	0.6217	5.3528	27.5587	0.0000	0.0000
ENSG00000132170	PPARG	0.6402	5.5378	23.0551	0.0000	0.0000
ENSG00000139515	PDX1	0.6714	0.1886	2.6065	0.1064	0.3237
ENSG00000134954	ETS1	0.7010	4.8578	20.7411	0.0000	0.0001
ENSG00000182759	MAFA	0.7297	1.1876	9.7487	0.0018	0.0177
ENSG00000138378	STAT4	0.7530	1.0413	5.5393	0.0186	0.1010
ENSG00000113916	BCL6	0.7550	5.5013	25.0300	0.0000	0.0000
ENSG00000115415	STAT1	0.7756	7.6784	45.7658	0.0000	0.0000
ENSG00000179388	EGR3	1.1668	1.8093	29.3749	0.0000	0.0000
ENSG00000179348	GATA2	1.1681	5.7637	24.0412	0.0000	0.0000
ENSG00000165556	CDX2	3.1288	5.0992	71.7325	0.0000	0.0000

[1]Possible transcription factor binding sites within the 2.5 kb promoter region of MIR100HG were predicted by the Match program (version 1.0).

[2]Candidate transcription factors differentially expressed between CC and CC-CR (fold change >1.5) were listed.

Extended Data Table 7

Clinic-pathological characteristics of metastatic CRC patients with paired specimen pre- and post-cetuximab treatment

No.[1]	Gender	Age	Primary site	Metastatic sites	Differentiation[2]	Cetuximab regimen[3]	Best response to cetuximab[4]	Site of specimen (pre)
1	M	56	left colon	lung, liver, peritoneal cavity	G2	FOLFOX4+cetuximab	PR	sigmoid
2	F	39	left colon	left ovary and adnexa, lung	G1	FOLFIRI+cetuximab	SD	sigmoid
3	M	52	right colon	lung, liver, peritoneal cavity	G1-G2	FOLFIRI+cetuximab	SD	Hepatic flexure of colon
4	M	52	left colon	liver	G2-G3	FOLFIRI+cetuximab	SD	left colon
5	F	55	left colon	liver	G1	FOLFOX4+cetuximab	PR	sigmoid
6	M	55	rectosigmoid	liver	G2	FOLFOX4+cetuximab	PR	rectum
7	F	48	left colon	lung, liver, peritoneal cavity	G2	mFOLFOX6+cetuximab	SD	sigmoid
8	F	81	transverse colon	liver, bone	G3	Cetuximab alone	SD	transverse colon
9	M	57	left colon	liver and lung	G2	FOLFOX4+cetuximab	PR	sigmoid
10	M	72	rectum	liver	G2	FOLFOX4+cetuximab	PR	rectum

[1] Cases 1, 2, 3 in Fig. 6d denote subject No. 2, 3, 5 in this table.

[2] G1, well-differentiated, G2, moderately differentiated, G3, poorly differentiated.

[3] Cetuximab 400 mg/m^2 initial dose followed by 250 mg/m^2 weekly thereafter with cetuximab dose intensity>90% were given to all subjects. Chemotherapy regimens: FOLFOX4 or mFOLFOX6, modified FOLFOX6 (5-fluorouracil, leucovorin, and oxaliplatin).

[4] PR, Partial Response; SD, Stable Disease.

Extended Data Table 8

KRAS, NRAS, and *BRAF* mutational status and *MET* amplification status in CRC patients with paired specimens obtained prior to cetuximab (Pre) and at time of tumor progression (Post)

No.[1]	Combined analysis of *KRAS, NRAS*, and *BRAF*[2]		*MET/CEP7* ratio[3]		miR-100/125b expression change (Post vs Pre)
	Pre	Post	Pre	Post	
1	WT	WT	<2	<2	Up
2	WT	*NRAS* c.182A>T p.Q61L	<2	<2	Up
3	WT	WT	n.d.	<2	Up
4	WT	*KRAS* c.34G>A p.G12S	<2	<2	Up
5	WT	WT	<2	<2	Up
6	WT	WT	<2	<2	Down
7	WT	WT	<2	<2	Down
8	WT	WT	n.d.	<2	Up
9	WT	WT	<2	<2	Down

No.[1]	Combined analysis of *KRAS*, *NRAS*, and *BRAF*[2]		*MET/CEP7* ratio[3]		miR-100/125b expression change (Post vs Pre)
	Pre	Post	Pre	Post	
10	WT	WT	n.d.	<2	Up

[1] Cases 1, 2, 3 in Fig. 6d denote Subjects No. 2, 3, 5 in this table.

[2] *NRAS* Q61L and *KRAS* G12S were identified in post-treatment specimens of Subject No. 2 and 4, respectively. In these two cases, DNA isolated from both the pre-and post-treatment samples was sequenced in parallel for *KRAS*, *NRAS*, and *BRAF*.

[3] WT, wild-type; n.d., not detected.

Extended Data Table 9
Primers used in the qRT-PCR assays for indicated genes

Gene Name	Forward Primer Sequence (5′ to 3′)	Reverse Primer Sequence (5′ to 3′)
ACTB	GGACTTCGAGCAAGAGATGG	AGCACTGTGTTGGCGTACAG
DKK1	AACAGCTATCCAAATGCAG	TCACAGGGGAGTTCCATAAA
DKK3	CTGGGAGCTAGAGCCTGATG	TCATACTCATCGGGGACCTC
CCND1	TTCAAATGTGTGCAGAAGGA	GGGATGGTCTCCTTCATCTT
KLF4	CGAACCCACACAGGTGAGAA	TACGGTAGTGCCTGGTCAGTTC
MYC	ACCAGAGAAACCTAACAGTGC	CTCTTTCATTTCGGCCAGTTC
NKD1	TGCCTCCTGAGAAGACTGAC	CATAGATGGTGTGCAGCAAG
PROX1	TCACCTTATTCGGGAAGTGC	GTACTGGTGACCCCATCGTT
S100A4	AACTAAAGGAGCTGCTGACCC	TGTTGCTGTCCAAGTTGCTC
CD44	TAGGAGAAGGTGTGGGCAGAA	GAGCTCACTGGGTTTCCTGTCTT
FOSL1	AGTCAGGAGCTGCAGTGGATGGT	TCAGTTCCTTCCTCCGGTTCCTGC
GATA6	TGCAATGCTTGTGGACTCTA	GTGGGGGAAGTATTTTTGCT
AXIN2	TACCGGAGGATGCTGAAGGC	CCACTGGCCGATTCTTCCTT

Extended Data Table 10
Primers used in the qRT-PCR assays for indicated miRNA or lncRNA

Name	Life Technologies ID	Category	Species
MIR100HG	Hs03680804_m1	TaqMan ® LncRNA Assay	homo sapiens
pri-mir-100	Hs03302731_pri	TaqMan ® Pri-miRNA Assay	homo sapiens
pri-mir-125b-1	Hs03303095_pri	TaqMan ® Pri-miRNA Assay	homo sapiens
pri-let-7a-2	Hs03302539_pri	TaqMan ® Pri-miRNA Assay	homo sapiens
miR-100	# 4427975 000437	TaqMan® MicroRNA Assays	homo sapiens
miR-125b	# 4427975 000449	TaqMan® MicroRNA Assays	homo sapiens
let-7a	# 4427975 000377	TaqMan® MicroRNA Assays	homo sapiens
ACTB	Hs01060665_g1	TaqMan ® gene expression Assay	homo sapiens
U6 snRNA	# 4427975 001973	TaqMan® microRNA Control Assays	homo sapiens

Extended Data Table 11

Primers used for Sanger sequencing of *KRAS*, *NRAS*, and *BRAF*

Locus	Primer Code	Prime Sequence (5′ to 3′)
KRAS Codon G12, G13	KRAS-exon2-F	GTTCTAATATAGTCACATTTTCA
	KRAS-exon2-R	TCTATTGTTGGATCATATTCG
KRAS Codon Q61	KRAS-exon3-F	TCTCCCTTCTCAGGATTC
	KRAS-exon3-R	ATTATTTATGGCAAATACACAAAG
KRAS Codon A146	KRAS-exon4-F	TTCTAGAACAGTAGACACAAAAC
	KRAS-exon4-R	GAGAGAAAAACTGATATATTAAATGAC
KRAS Codon K117	KRAS-exon4-2F	CTTTCCCAGAGAACAAATTAAAAG
	KRAS-exon4-2R	TCAATAAAAGGAATTCCATAACTTCT
NRAS Codon G12, G13	NRAS-exon2-F	CTGATTACTGGTTTCCAACAG
	NRAS-exon2-R	CCTCTATGGTGGGATCATATTC
NRAS Codon Q61	NRAS-exon3-F	CCCCAGGATTCTTACAGAAAA
	NRAS-exon3-R	TTGATGGCAAATACACAGAG
BRAF Codon G465, G468, Y472	BRAF-exon11-2F	GGGACTCGAGTGATGA
	BRAF-exon11-2R	AAAAGTTGTTAAACATATCCTATT
BRAF Codon D593, F594, L596, L597, T598, V600, K601	BRAF-exon15-2F	ATGAGATCTACTGTTTTCCTTTACT
	BRAF-exon15-2R	CCTCAATTCTTACCATCCACA

Acknowledgments

We acknowledge the support of the Vanderbilt University Cell Imaging, Chemical Biology Synthesis, and Flow Cytometry Shared Resources. We thank J. Higginbotham for help with flow cytometry, W. Fry for help with plasmid construction, and E. Poulin and N. Markham for critical editing of the manuscript. We thank X. Wang and Y. Nie (Xijing Hospital of Digestive Diseases) for providing clinically annotated samples. This work was supported by National Cancer Institute (NCI) R01 CA046413, R35 CA197570, and P50 CA095103 GI Specialized Programs of Research Excellence to R.J.C., Natural Science Foundation of China (NSFC) 81430072 and National Key R&D Program 2016YFC1303200 to D.F. and Emmy Noether-Programme of the German Research Foundation KL 2374/2-1 to J.H.K.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7. [PubMed: 26742998]
2. Bertotti A, et al. The genomic landscape of response to EGFR blockade in colorectal cancer. Nature. 2015; 526:263. [PubMed: 26416732]
3. Stintzing S, et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab for metastatic colorectal cancer (FIRE-3): a post-hoc analysis of tumour dynamics in the final RAS wild-type subgroup of this randomised open-label phase 3 trial. Lancet Oncol. 2016; 17:1426. [PubMed: 27575024]
4. Misale S, Di Nicolantonio F, Sartore-Bianchi A, Siena S, Bardelli A. Resistance to anti-EGFR therapy in colorectal cancer: from heterogeneity to convergent evolution. Cancer Discov. 2014; 4:1269. [PubMed: 25293556]
5. Arena S, et al. Emergence of Multiple EGFR Extracellular Mutations during Cetuximab Treatment in Colorectal Cancer. Clin Cancer Res. 2015; 21:2157. [PubMed: 25623215]

6. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. Cell. 2014; 157:77. [PubMed: 24679528]
7. Adams BD, Parsons C, Walker L, Zhang WC, Slack FJ. Targeting noncoding RNAs in disease. J Clin Invest. 2017; 127:761. [PubMed: 28248199]
8. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. Genome Res. 2004; 14:1902. [PubMed: 15364901]
9. Dhir A, Dhir S, Proudfoot NJ, Jopling CL. Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs. Nat Struct Mol Biol. 2015; 22:319. [PubMed: 25730776]
10. Keniry A, et al. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat Cell Biol. 2012; 14:659. [PubMed: 22684254]
11. Dews M, et al. The myc-miR-17~92 axis blunts TGF{beta} signaling and production of multiple TGF{beta}-dependent antiangiogenic factors. Cancer Res. 2010; 70:8233. [PubMed: 20940405]
12. Emmrich S, et al. miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFbeta and Wnt signaling. Genes Dev. 2014; 28:858. [PubMed: 24736844]
13. Emmrich S, et al. LincRNAs MONC and MIR100HG act as oncogenes in acute megakaryoblastic leukemia. Mol Cancer. 2014; 13:171. [PubMed: 25027842]
14. Li C, et al. Excess PLAC8 promotes an unconventional ERK2-dependent EMT in colon cancer. J Clin Invest. 2014; 124:2172. [PubMed: 24691442]
15. Li C, et al. Three-dimensional culture system identifies a new mode of cetuximab resistance and disease-relevant genes in colorectal cancer. Proc Natl Acad Sci U S A. 2017; 114:E2852. [PubMed: 28320945]
16. Jhawer M, et al. PIK3CA mutation/PTEN expression status predicts response of colon cancer cells to the epidermal growth factor receptor inhibitor cetuximab. Cancer Res. 2008; 68:1953. [PubMed: 18339877]
17. Medico E, et al. The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. Nat Commun. 2015; 6:7002. [PubMed: 25926053]
18. Augoff K, McCue B, Plow EF, Sossey-Alaoui K. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. Mol Cancer. 2012; 11:5. [PubMed: 22289355]
19. Schell MJ, et al. A multigene mutation classification of 468 colorectal cancers reveals a prognostic role for APC. Nat Commun. 2016; 7:11743. [PubMed: 27302369]
20. Dry JR, et al. Transcriptional pathway signatures predict MEK addiction and response to selumetinib (AZD6244). Cancer Res. 2010; 70:2264. [PubMed: 20215513]
21. Hausser J, Zavolan M. Identification and consequences of miRNA-target interactions--beyond repression of gene expression. Nat Rev Genet. 2014; 15:599. [PubMed: 25022902]
22. Meerbrey KL, et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc Natl Acad Sci U S A. 2011; 108:3665. [PubMed: 21307310]
23. Huang SM, et al. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature. 2009; 461:614. [PubMed: 19759537]
24. Emami KH, et al. A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. Proc Natl Acad Sci U S A. 2004; 101:12682. [PubMed: 15314234]
25. Kel AE, et al. MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res. 2003; 31:3576. [PubMed: 12824369]
26. Haveri H, et al. Transcription factors GATA-4 and GATA-6 in normal and neoplastic human gastrointestinal mucosa. BMC Gastroenterol. 2008; 8:9. [PubMed: 18405344]
27. Whissell G, et al. The transcription factor GATA6 enables self-renewal of colon adenoma stem cells by repressing BMP gene expression. Nat Cell Biol. 2014; 16:695. [PubMed: 24952462]
28. Martinelli P, et al. GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. Gut. 2016

29. Aronson BE, Stapleton KA, Krasinski SD. Role of GATA factors in development, differentiation, and homeostasis of the small intestinal epithelium. Am J Physiol Gastrointest Liver Physiol. 2014; 306:G474. [PubMed: 24436352]
30. Kikuchi K, et al. Transcripts of unknown function in multiple-signaling pathways involved in human stem cell differentiation. Nucleic Acids Res. 2009; 37:4987. [PubMed: 19531736]
31. Shang C, et al. Characterization of long non-coding RNA expression profiles in lymph node metastasis of early-stage cervical cancer. Oncol Rep. 2016
32. Chen D, et al. miR-100 induces epithelial-mesenchymal transition but suppresses tumorigenesis, migration and invasion. Plos Genet. 2014; 10:e1004177. [PubMed: 24586203]
33. Ueda T, et al. Relation between microRNA expression and progression and prognosis of gastric cancer: a microRNA expression analysis. Lancet Oncol. 2010; 11:136. [PubMed: 20022810]
34. Cruciat CM, Niehrs C. Secreted and transmembrane wnt inhibitors and activators. Cold Spring Harb Perspect Biol. 2013; 5:a15081.
35. MacDonald BT, Tamai K, He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell. 2009; 17:9. [PubMed: 19619488]
36. Bafico A, Liu G, Yaniv A, Gazit A, Aaronson SA. Novel mechanism of Wnt signalling inhibition mediated by Dickkopf-1 interaction with LRP6/Arrow. Nat Cell Biol. 2001; 3:683. [PubMed: 11433302]
37. Nakamura RE, Hackam AS. Analysis of Dickkopf3 interactions with Wnt signaling receptors. Growth Factors. 2010; 28:232. [PubMed: 20370576]
38. de Lau W, Peng WC, Gros P, Clevers H. The R-spondin/Lgr5/Rnf43 module: regulator of Wnt signal strength. Genes Dev. 2014; 28:305. [PubMed: 24532711]
39. Jiang X, Charlat O, Zamponi R, Yang Y, Cong F. Dishevelled promotes Wnt receptor degradation through recruitment of ZNRF3/RNF43 E3 ubiquitin ligases. Mol Cell. 2015; 58:522. [PubMed: 25891077]
40. Bond CE, et al. RNF43 and ZNRF3 are commonly altered in serrated pathway colorectal tumorigenesis. Oncotarget. 2016
41. Schneikert J, et al. Functional comparison of human adenomatous polyposis coli (APC) and APC-like in targeting beta-catenin for degradation. Plos One. 2013; 8:e68072. [PubMed: 23840886]
42. Croy HE, et al. The Poly(ADP-ribose) Polymerase Enzyme Tankyrase Antagonizes Activity of the beta-Catenin Destruction Complex through ADP-ribosylation of Axin and APC2. J Biol Chem. 2016; 291:12747. [PubMed: 27068743]
43. Yu X, et al. CXCL12/CXCR4 axis induced miR-125b promotes invasion and confers 5-fluorouracil resistance through enhancing autophagy in colorectal cancer. Sci Rep. 2017; 7:42226. [PubMed: 28176874]
44. Zhang L, Ge Y, Fuchs E. miR-125b can enhance skin tumor initiation and promote malignant progression by repressing differentiation and prolonging cell survival. Genes Dev. 2014; 28:2532. [PubMed: 25403182]
45. Paul I, Bhattacharya S, Chatterjee A, Ghosh MK. Current Understanding on EGFR and Wnt/beta-Catenin Signaling in Glioma and Their Possible Crosstalk. Genes Cancer. 2013; 4:427. [PubMed: 24386505]
46. Hu T, Li C. Convergence between Wnt-beta-catenin and EGFR signaling in cancer. Mol Cancer. 2010; 9:236. [PubMed: 20828404]
47. Horst D, et al. Differential WNT activity in colorectal cancer confers limited tumorigenic potential and is regulated by MAPK signaling. Cancer Res. 2012; 72:1547. [PubMed: 22318865]
48. Civenni G, Holbro T, Hynes NE. Wnt1 and Wnt5a induce cyclin D1 expression through ErbB1 transactivation in HC11 mammary epithelial cells. Embo Rep. 2003; 4:166. [PubMed: 12612606]
49. Casas-Selves M, et al. Tankyrase and the canonical Wnt pathway protect lung cancer cells from EGFR inhibition. Cancer Res. 2012; 72:4154. [PubMed: 22738915]
50. Nakayama S, et al. beta-catenin contributes to lung tumor development induced by EGFR mutations. Cancer Res. 2014; 74:5891. [PubMed: 25164010]
51. Tan X, et al. Epidermal growth factor receptor: a novel target of the Wnt/beta-catenin pathway in liver. Gastroenterology. 2005; 129:285. [PubMed: 16012954]

52. Zhong Y, et al. GATA6 activates Wnt signaling in pancreatic cancer by negatively regulating the Wnt antagonist Dickkopf-1. Plos One. 2011; 6:e22129. [PubMed: 21811562]
53. Martinelli P, et al. GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. Gut. 2016
54. Martinelli P, et al. The acinar regulator Gata6 suppresses KrasG12V-driven pancreatic tumorigenesis in mice. Gut. 2016; 65:476. [PubMed: 25596178]
55. Tsuji S, et al. The miR-363-GATA6-Lgr5 pathway is critical for colorectal tumourigenesis. Nat Commun. 2014; 5:3150. [PubMed: 24452072]
56. Kawasaki Y, et al. REG4 is a transcriptional target of GATA6 and is essential for colorectal tumorigenesis. Sci Rep. 2015; 5:14291. [PubMed: 26387746]
57. Fodde R, Brabletz T. Wnt/beta-catenin signaling in cancer stemness and malignant behavior. Curr Opin Cell Biol. 2007; 19:150. [PubMed: 17306971]
58. Schneikert J, Grohmann A, Behrens J. Truncated APC regulates the transcriptional activity of beta-catenin in a cell cycle dependent manner. Hum Mol Genet. 2007; 16:199. [PubMed: 17189293]
59. Voloshanenko O, et al. Wnt secretion is required to maintain high levels of Wnt activity in colon cancer cells. Nat Commun. 2013; 4:2610. [PubMed: 24162018]
60. Kim D, et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36. [PubMed: 23618408]
61. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26:139. [PubMed: 19910308]
62. Chen CJ, et al. ncPRO-seq: a tool for annotation and profiling of ncRNAs in sRNA-seq data. Bioinformatics. 2012; 28:3147. [PubMed: 23044543]
63. Cingolani P, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012; 6:80. [PubMed: 22728672]
64. Saunders CT, et al. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics. 2012; 28:1811. [PubMed: 22581179]
65. Ebert MS, Sharp PA. MicroRNA sponges: progress and possibilities. RNA. 2010; 16:2043. [PubMed: 20855538]
66. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods. 2007; 4:721. [PubMed: 17694064]
67. Meerbrey KL, et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proc Natl Acad Sci U S A. 2011; 108:3665. [PubMed: 21307310]
68. Zhao XD, et al. MicroRNA-7/NF-kappaB signaling regulatory feedback circuit regulates gastric carcinogenesis. J Cell Biol. 2015; 210:613. [PubMed: 26261179]
69. de Planell-Saguer M, Rodicio MC, Mourelatos Z. Rapid in situ codetection of noncoding RNAs and proteins in cells and formalin-fixed paraffin-embedded tissue sections without protease treatment. Nat Protoc. 2010; 5:1061. [PubMed: 20539282]
70. Schell MJ, et al. A Composite Gene Expression Signature Optimizes Prediction of Colorectal Cancer Metastasis and Outcome. Clin Cancer Res. 2016; 22:734. [PubMed: 26446941]
71. Herbst A, et al. Comprehensive analysis of beta-catenin target genes in colorectal carcinoma cell lines with deregulated Wnt/beta-catenin signaling. BMC Genomics. 2014; 15:74. [PubMed: 24467841]
72. Jorissen RN, et al. Metastasis-Associated Gene Expression Changes Predict Poor Outcomes in Patients with Dukes Stage B and C Colorectal Cancer. Clin Cancer Res. 2009; 15:7642. [PubMed: 19996206]
73. Marisa L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. Plos Med. 2013; 10:e1001453. [PubMed: 23700391]

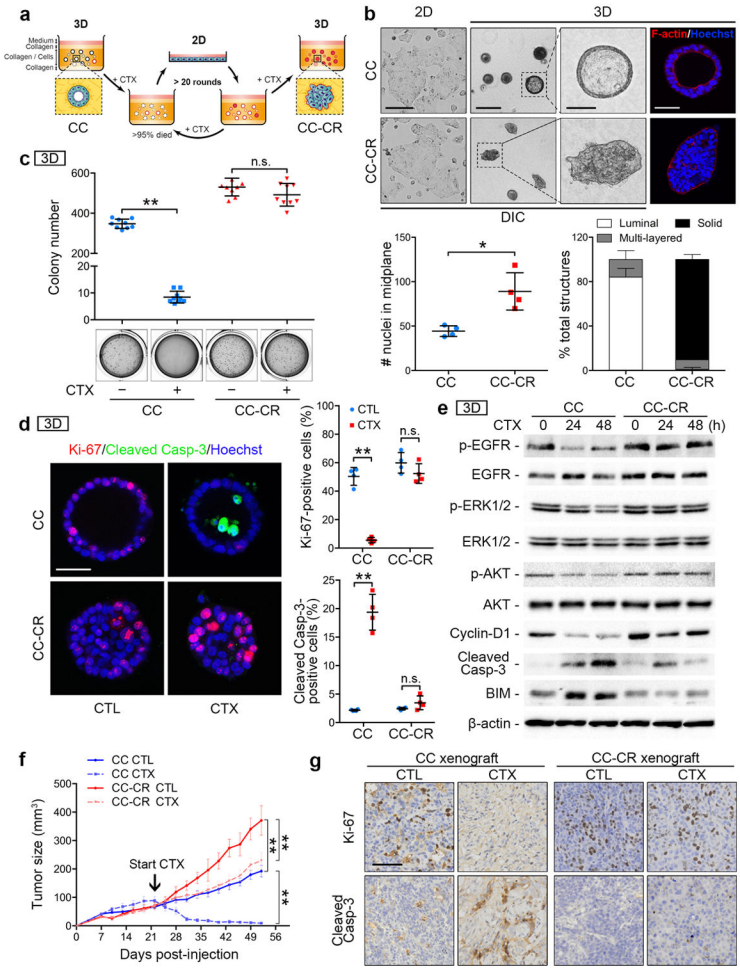


Figure 1. Characterization of cetuximab-resistant CC (CC-CR) in 3D

(**a**) Schematic of experimental approach to establish cetuximab (CTX)-resistant cells in 3D. In the presence of CTX (3 μg/ml) in 3D type-1 collagen culture, greater than 95% of CC colonies die. Residual colonies were isolated and iteratively passaged in 2D and 3D in the continued presence of CTX over approximately 4 months. These colonies were designated CC-CR. (**b**) Top: differential interference contrast (DIC) and confocal images of representative CC and CC-CR in 2D and 3D. F-actin was stained with phalloidin (red). Scale bars: 400, 1000, 200, 50 μm, respectively (from left to right). Bottom: left, number of nuclei in the midplane of each colony; right, the morphology of colonies was divided into those with luminal, multi-layered, or solid morphology. n=4 independent experiments, **P*<0.05 by Student's *t* test. (**c**) CC and CC-CR were cultured in 3D in the presence or absence of CTX (3 μg/ml) and colonies were counted after 18 days. n=3 independent experiments performed in triplicate. ***P*<0.01 by Student's *t* test. (**d**) CC and CC-CR cells were cultured in 3D for 12 days and treated with CTX (10 μg/ml) for 24 h. Ki-67 (red) and Cleaved Caspase-3 (Cleaved Casp-3, green) staining were imaged by confocal microscopy. Representative of 4 independent experiments. Scale bar, 50 μm. Quantification is shown on the right (n=4). ***P*<0.01 by Student's *t* test. (**e**) Immunoblots of 3D cell lysates from CC and CC-CR treated with CTX (10 μg/ml) for indicated time. β-actin served as the loading control. A

representative blot from 3 independent experiments is shown. (**f**) Nude mice (n=8) bearing subcutaneous tumors were treated with control saline or CTX at a dose of 1 mg/mouse, intraperitoneal (i.p.) injection, every 3 days. Tumor volumes were measured every 3 days using calipers. $^{**}P<0.01$ by repeated-measures ANOVA test followed by LSD post-hoc test. (**g**) Representative immunohistochemical images of Ki-67 and Cleaved Casp-3 from CC and CC-CR xenografts before and after CTX treatment. Scale bar: 50 μm. Data represent mean ± s.d. in b-d and f. n.s., not significant.

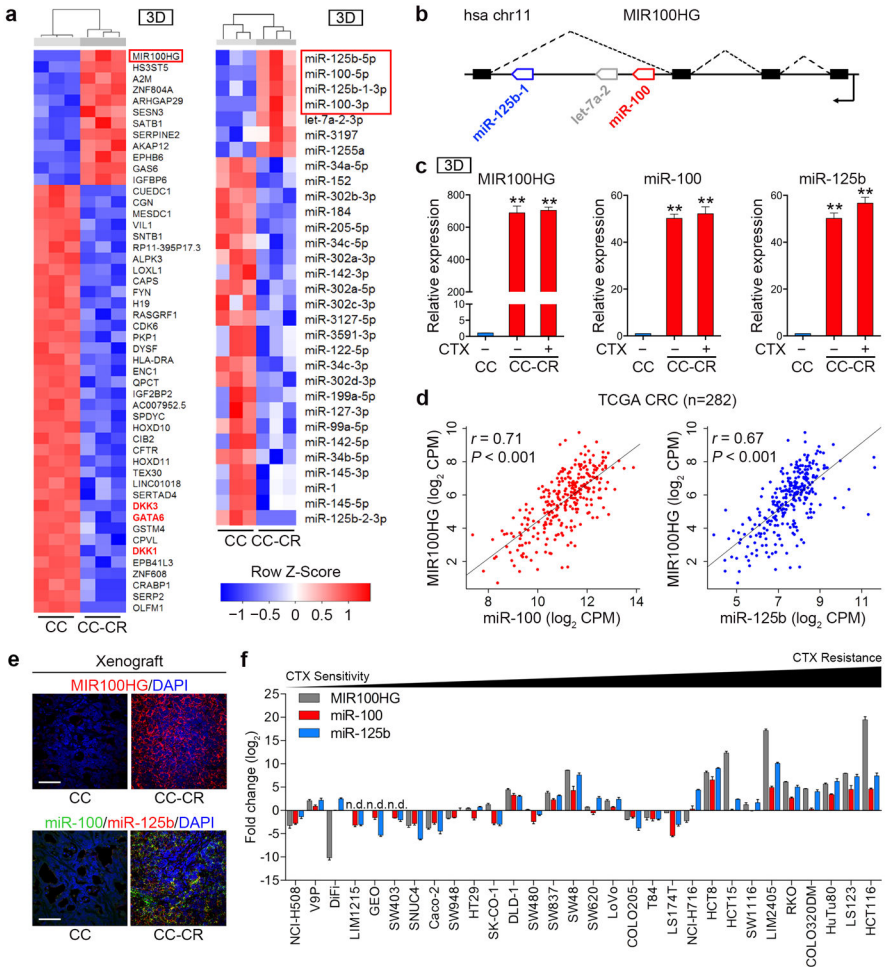


Figure 2. Transcriptome profiling of CC and CC-CR in 3D

(**a**) Left, heatmap of top 50 differentially expressed transcripts in CC-CR versus CC from 3 independent 3D culture experiments. Gene expression values are gene-wise z-transformed and are colored red for high abundance and blue for low abundance, as indicated in the scale bar. Right, miRNA heatmap showing miRNAs altered (>2-fold and FDR<0.01) in CC-CR versus CC. (**b**) Genomic organization of lncRNA MIR100HG, host gene of miR-100/let-7a-2/miR-125b-1 cluster, on human chromosome 11 (hsa chr11). (**c**) qRT-PCR showing upregulation of lncRNA MIR100HG, miR-100 and miR-125b in CC-CR compared to CC in 3D. In CC-CR, cells were treated with CTX (CTX+, 3 μg/ml) or normal culture medium (CTX-) for 14 consecutive days in 3D. ACTB or U6 snRNA served as the internal control, respectively. n=3 independent experiments performed in triplicate. ** P<0.01 by one-way ANOVA followed by Dunnett's test. (**d**) Scatter plots of MIR100HG versus miR-100 or miR-125b expression in TCGA CRC data repository. Pearson correlation coefficients (r) and P values are shown. (**e**) RNA FISH showing high MIR100HG (red) expression in CC-CR mouse tumor xenografts compared to CC xenografts. Concomitantly, high miR-100 (green) and miR-125b (red) signals were observed in CC-CR tumors; the yellow fluorescent signal indicates co-expression of miR-100 and miR-125b. Scale bars, 50 μm. (**f**) qRT-PCR analysis of MIR100HG, miR-100 and miR-125b expression levels among a panel of 30 CRC cell

lines ranked by their responsiveness to cetuximab (see Extended Data Table 3). ACTB or U6 snRNA served as the internal control. Fold changes were normalized to CC. n.d., not detected. n=3 independent experiments performed in triplicate. Data represent mean ± s.d. in c and f.

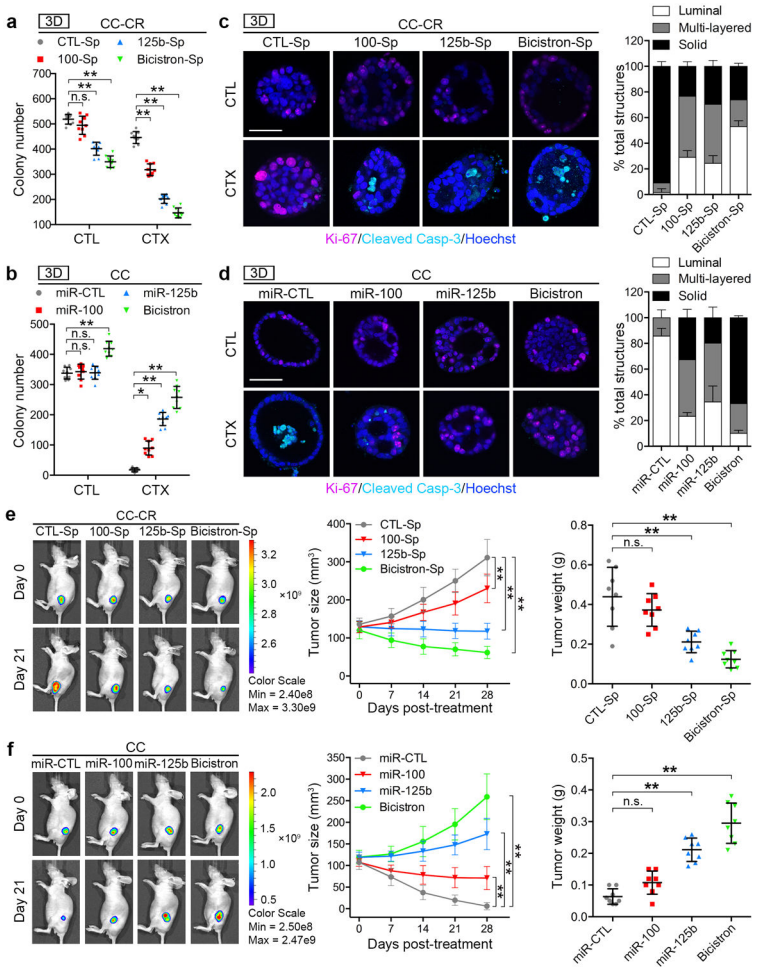


Figure 3. Cooperativity of miR-100 and miR-125b in CTX resistance

(**a, b**) Indicated cells were grown in 3D in normal medium (CTL) or treated with CTX (3 μg/ml) in 3D. The resultant colonies were counted after 18 days. Sponge (Sp). n=3 independent experiments performed in triplicate. **P*<0.05, ***P*<0.01 by one-way ANOVA followed by Dunnett's test compared with CTL-Sp or miR-CTL. (**c, d**) Left: indicated cells were cultured in 3D for 12 days and CTX (10 μg/ml) was added for 24 h before cells were fixed, stained for Cleaved Casp-3 (cyan) and Ki-67 (magenta). Scale bars, 50 μm. Right: quantification of the morphological changes among indicated cell lines. n=4 independent experiments. (**e, f**) Left: indicated cells were injected subcutaneously into nude mice (n=8). After tumor size reached approximately 100 mm^3, the mice received CTX treatment (1 mg/mouse, i.p. injection every 3 days). Representative fluorescent images of GFP signals captured from subcutaneous tumors are shown. Middle: growth curve of tumors in nude mice (n=8) injected with cells as indicated. ***P*<0.01 by repeated-measures ANOVA test followed by Dunnett's test. Right: tumors (n=8) were isolated on day 28 after treatment and tumor weight was calculated. ***P*<0.01 by one-way ANOVA followed by Dunnett's test compared with CTL-Sp or miR-CTL. Data represent mean ± s.d. n.s., not significant.

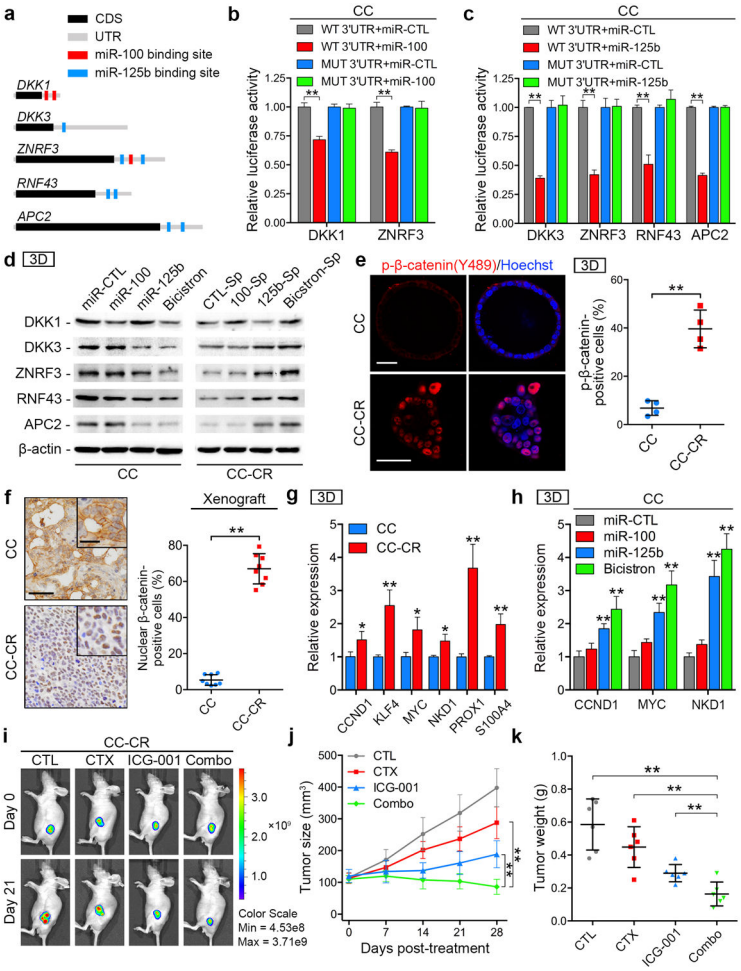


Figure 4. miR-100 and miR-125b augment Wnt signaling by repressing multiple Wnt negative regulators

(**a**) Predicted miR-100 (red) and miR-125b (blue) binding sites in 3′ untranslated regions (3′ UTRs) of human DKK1, DKK3, ZNRF3, RNF43, and APC2. CDS, coding sequence. (**b, c**) Dual luciferase assays of candidates predicted to be regulated by miR-100 or miR-125b. *Renilla* luciferase activity was normalized to firefly activity and presented as relative luciferase activity. n=2 independent experiments. ***P*<0.01 by Student's *t* test. (**d**) Immunoblots of indicated proteins in stable miRNA-transduced CC and sponge (Sp)-transduced CC-CR. Representative of 3 independent experiments. (**e**) Immunofluorescence of p-β-catenin (Y489). Scale bars, 50 μm. Right, quantification of 4 independent experiments. ***P*<0.01 by Student's *t* test. (**f**) Representative IHC of β-catenin in CC and CC-CR xenografts (n=8). Scale bars: 50 μm (main); 20 μm (inset). Quantification of nuclear β-catenin-positive cells is shown. ***P*<0.01 by Student's *t* test. (**g**) qRT-PCR analysis of Wnt target genes in CC and CC-CR cells. n=3 independent experiments performed in triplicate. **P*<0.05, ***P*<0.01 by Student's *t* test. (**h**) qRT-PCR analysis of Wnt target genes in the indicated stable miRNA-transduced CC cells. n=3 independent experiments performed in triplicate. ***P*<0.01 by one-way ANOVA followed by Dunnett's test. (**i**) CC-CR were injected subcutaneously into nude mice (n=6). When tumor size reached around 100 mm^3, mice were treated with control saline, or CTX (1 mg/mouse, i.p. injection every 3 days)

and/or ICG-001 (150 mg/kg i.p. injection daily). Representative *in vivo* fluorescent images are shown. (**j**) Growth curve of tumors in nude mice (n=6) treated with different compounds. $^{**}P<0.01$ by repeated-measures ANOVA test followed by LSD post-hoc test. (**k**) Tumors (n=6) were isolated on day 28 after treatment and tumor weight was measured. $^{**}P<0.01$ by one-way ANOVA followed by LSD post-hoc test. Data represent mean ± s.d.

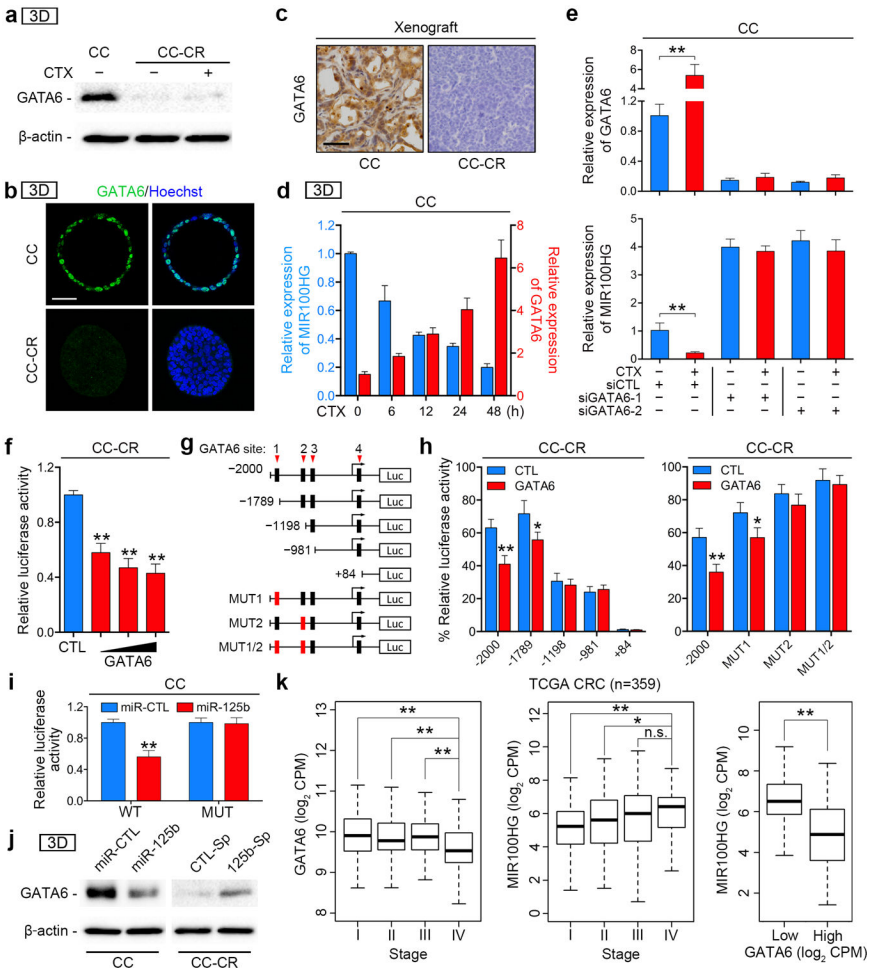


Figure 5. GATA6 transcriptionally represses MIR100HG and is targeted by miR-125b in a double-negative feedback loop

(**a**) Immunoblot of GATA6 in CC and CC-CR cells cultured in 3D. In CC-CR, cells were treated with CTX (CTX+, 3 μg/ml) or normal culture medium (CTX-) for consecutive 14 days in 3D before protein extraction. Representative of 3 independent experiments. (**b**) Immunofluorescence of GATA6 (green) and nuclei (blue). Scale bar, 50 μm. (**c**) Representative IHC of GATA6 in CC and CC-CR xenografts (n=8). (**d**) qRT-PCR analysis of MIR100HG and GATA6 expression at indicated time points following CTX treatment (10 μg/ml) in CC cultured in 3D. n=3 independent experiments. (e) CC cells were transfected with two independent siRNAs against GATA6 or control (siCTL), treated with CTX (10 μg/ml) and subjected to qRT-PCR analysis. n=2 independent experiments performed in triplicate. ***P*<0.01 by Student's *t* test. (**f**) Luciferase reporter assays were performed by co-transfection of pGL3-MIR100HG promoter luciferase reporter with increasing concentrations of pcDNA3.1-GATA6 plasmid or empty vector control (CTL), along with a *Renilla* luciferase reporter. Luciferase activity was measured 36 h post-transfection and normalized to *Renilla* values. n=3 independent experiments performed in triplicate. ***P*<0.01 by one-way ANOVA followed by Dunnett's test. (**g**) A schematic representation of consecutive deletion and mutation constructs spanning the -2000~+500 region of MIR100HG promoter. The putative GATA6-binding sites within MIR100HG promoter are

shown in black boxes. (**h**) The luciferase vector pGL3 driven by either wild-type, deletion or mutant (MUT) promoter was transfected in CC-CR, and luciferase activity was measured. n=3 independent experiments. **P*<0.05, ***P*<0.01 by Student's *t* test. (**i**) Luciferase reporter analysis of a wild-type (WT) or mutant (MUT) GATA6 3′ UTR activity upon addition of either synthetic miR-125b or a negative control miR-CTL. ***P*<0.01 by Student's *t* test. (**j**) Immunoblots of GATA6 in stable miR-125b-transduced CC and 125b-Sp-transduced CC-CR. Representative of 3 independent experiments. (**k**) Box plots showing expression of GATA6 (left) and MIR100HG (middle) by stage from the TCGA CRC data repository. Right panel depicts MIR100HG expression in the lower (<25%) and the higher (>75%) quartile of GATA6 expression. **P*<0.05, ***P*<0.01 by Mann-Whitney *U* test. n.s., not significant. Data represent mean ± s.d. in d-f, h, and i.

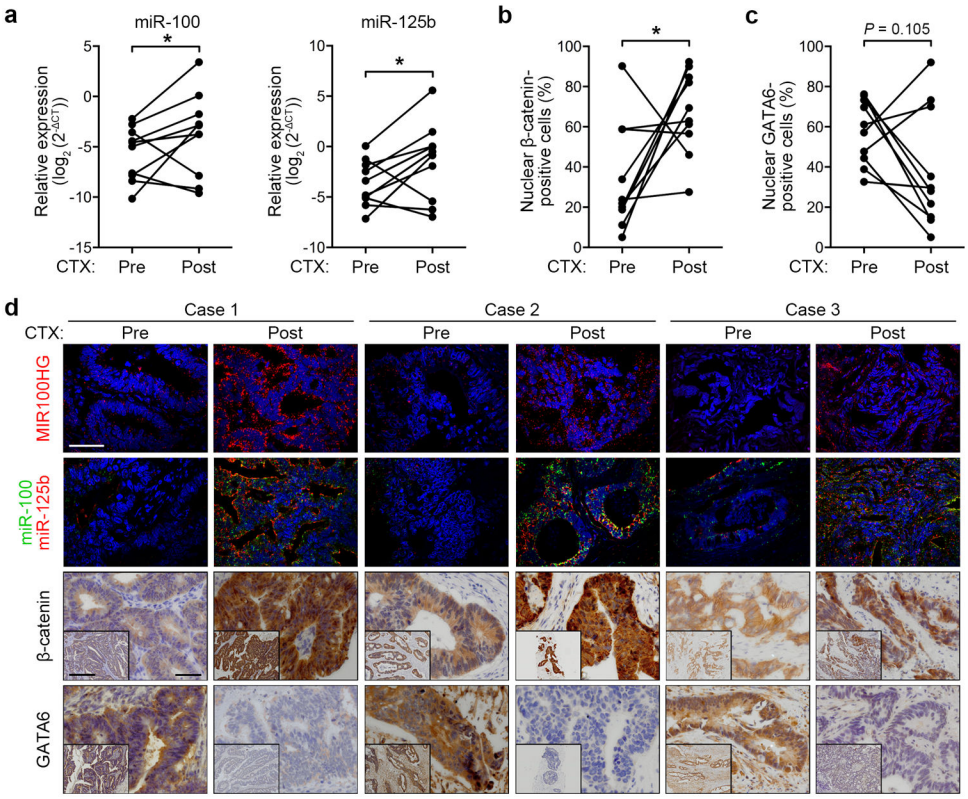


Figure 6. Increased MIR100HG and miR-100/125b are found in CRC patient specimens at time of progression on cetuximab

(**a**) qRT-PCR of miR-100 and miR-125b levels in 10 pairs of matched human CRC specimens pre- and post-cetuximab resistance. Each symbol represents mean value of an individual patient. **P*<0.05 by Wilcoxon matched-pairs signed rank test. (**b, c**) Frequency of nuclear β-catenin-positive cells (**b**) and GATA6-positive cells (**c**) in 10 pairs of matched human CRC specimens pre- and post-cetuximab resistance. **P*<0.05 by Wilcoxon matched-pairs signed rank test. (**d**) Representative FISH images of MIR100HG, miR-100/125b and corresponding IHC images of β-catenin and GATA6 staining in representative three paired human CRC specimens obtained pre- and post-cetuximab resistance. Scale bars, 50 μm (main); 500 μm (inset).