



UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets

This is the author's manuscript
Original Citation:
Availability:
This version is available http://hdl.handle.net/2318/1509370 since 2020-11-06T22:40:28Z
Published version:
DOI:10.1038/ncomms8002
Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on: Questa è la versione dell'autore dell'opera: Title: The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets

Authors: E. Medico, M. Russo, G. Picco, C. Cancelliere, E. Valtorta, G. Corti, M. Buscarino, C. Isella, S. Lamba, B. Martinoglio, S. Veronese, S. Siena, A. Sartore-Bianchi, M. Beccuti, M. Mottolese, M. Linnebacher, F. Cordero, F. Di Nicolantonio and A. Bardelli

Year: 2015

Journal: Nature communications

Volume: 6

DOI: 10.1038/ncomms8002

The definitive version is available at: La versione definitiva è disponibile alla URL: http://dx.doi.org/10.1038/ncomms8002

The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets

Enzo Medico^{1,2,*,§}, Mariangela Russo^{1,2,3,*}, Gabriele Picco^{1,2,*}, Carlotta Cancelliere^{2,*}, Emanuele Valtorta⁴, Giorgio Corti², Michela Buscarino², Claudio Isella¹, Simona Lamba^{1,2}, Barbara Martinoglio², Silvio Veronese⁴, Salvatore Siena⁴, Andrea Sarote-Bianchi⁴, Marco Beccuti⁵, Marcella Mottolese⁶, Michael Linnebacher⁷, Francesca Cordero⁵, Federica Di Nicolantonio^{1,2, §}, Alberto Bardelli^{1,2, §}

¹ University of Torino, Department of Oncology, SP 142, Km 3.95, 10060 Candiolo, Torino, Italy; ² Candiolo Cancer Institute – FPO, IRCCS, Candiolo, Torino, Italy; ³ FIRC Institute of Molecular Oncology (IFOM), Milano, Italy; ⁴ Department of Hematology and Oncology, Niguarda Cancer Center, Ospedale Niguarda Ca' Granda, Milan, Italy; ⁵ University of Torino, Department of Computer Science, Torino, Italy; ⁶ S.C. Anatomia Patologica, Istituto Nazionale Tumori Regina Elena –Rome, Italy; ⁷ Division of Molecular Oncology and Immunotherapy, University of Rostock, D-18057 Rostock, Germany

* co-first authors

§ co-senior authors

Correspondence to: Federica Di Nicolantonio (<u>federica.dinicolantonio@unito.it</u>) and Alberto Bardelli (<u>alberto.bardelli@unito.it</u>)

Short title: A colorectal cancer cell encyclopedia

Abstract

The development of molecularly targeted anticancer agents relies on large panels of tumor-specific preclinical models closely recapitulating the molecular heterogeneity observed in patients. Here we describe the mutational and gene expression analyses of 151 colorectal cancer (CRC) cell lines. We find that the whole spectrum of CRC molecular and transcriptional subtypes, previously defined in patients, is represented in this cell line compendium. Transcriptional outlier analysis identifies RAS/BRAF wild type cells, resistant to EGFR blockade, functionally and pharmacologically addicted to kinase genes including ALK, FGFR2, NTRK1/2 and RET. The same genes are present as expression outliers in CRC patient samples. Genomic rearrangements (translocations) involving the ALK and NTRK1 genes are associated with overexpression of the corresponding proteins in CRC specimens. The approach described here can be used to pinpoint CRCs with exquisite dependencies to individual kinases for which clinically approved drugs are already available.

Introduction

Knowledge of tumor biology and the development of new anticancer agents depend on robust preclinical model systems that reflect the genomic heterogeneity of human cancers and for which detailed genetic and pharmacologic annotations are available. Cell lines represent a mainstay to functionalize molecular data as they allow experimental manipulation, global and detailed mechanistic studies, and high-throughput applications^{1, 2, 3, 4}. Virtually all commonly used cancer cells were continuously grown *in vitro* for years, and decades have often passed since they were originally derived from patients. The *in*

vitro culture conditions used to propagate cancer cells are far from the histological landscape in which they originated. These considerations are often cited to question the relevance of cell lines as cancer models. We further observe that, for a given cancer type, only few cell lines (<10) are employed in most preclinical studies. The evidence that cancer patients have heterogeneous genetic features implies that a large number of lineage-specific cell lines is needed to capture the diversity observed in the clinic. Furthermore, cells commonly used to model a tumor type may not represent the patients that they intend to recapitulate. For instance, colorectal cancer (CRC) cell lines ordinarily used in preclinical studies often display microsatellite instability (MSI). MSI colorectal tumors are found in 10-15% patients ⁵, but they show an indolent clinical behavior and are less prevalent in more advanced stages of the disease, accounting for less than 5% in the metastatic setting ^{6, 7}. Consequently, MSI cell lines do not properly recapitulate the clinical setting in which targeted agents are most commonly administered to CRC patients.

In addition to 'genetic' driven subtypes, tumor lineages can be subdivided based on their transcriptional profiles. Transcriptional profiling has been recently used to identify distinct CRC subtypes ^{8, 9, 10, 11, 12}. The subtypes are associated with biological and clinical features such as cell of origin, microsatellite instability, prognosis and response to treatments. Whether the transcriptional subtypes are maintained in CRC cell models and how this knowledge can be used to discover novel pharmacogenetic relationships has not been explored. To address these challenges we assembled and annotated a comprehensive collection of 151 human CRC cells.

Please add a short paragraph (2-3 sentences) to summarize the main findings of your manuscript. Those must be in the present tense.

We find that the molecular heterogeneity (oncogenic mutations and transcriptional subtypes) previously defined in CRC patients, is maintained in colorectal cancer cells. Individual lines can be stratified as responders or non-responders to EGFR blockade on the basis of clinically validated biomarkers. Transcriptional outlier analysis identifies CRC cells, resistant to EGFR blockade, pharmacologically addicted to kinase genes including *ALK, FGFR2, NTRK1/2* and *RET*.

Results

Genetic and transcriptional profiling of CRC cells

A collection of 152 CRC cell lines was initially genotyped at ten different microsatellite loci (STR profile) to unequivocally define their genetic identities (Supplementary Data 1). This analysis revealed that a few cell lines, previously thought to be unrelated, were derived from the same individual (Supplementary Data 1). Global mRNA expression profiling was carried out on the entire cell bank and TiGER (tissue-specific gene expression database)¹³ was exploited to perform a "tissue of origin" analysis (Supplementary Fig. 1a). The analysis revealed a probable non-intestinal origin for one line, COLO741, found to express skin-specific genes, including Tyrosinase (Supplementary Fig. 1b,c). This, together with the lack of expression of the neuroendocrine markers CHGA, SYP, ENO2, ACPP and NCAM1, suggested that these cells are not derived from a CRC nor from a neuroendocrine tumor, but rather from a melanoma. For this reason, COLO741 cells were withdrawn from further analyses. HuTu80 cells displayed a small intestine-like signature (Supplementary Fig. 1d) and were included in subsequent analyses due to their intestinal origin. Gene expression-based hierarchical clustering revealed that cell lines with identical genetic background consistently clustered together (Fig. 1), indicating that the genotype maintained strong control over the transcriptome. The only exception was the COGA5/COGA5L pair; these lines were derived respectively from the primary tumor and a lymph node metastasis of the same patient ¹⁴.

The 151 CRC cell line collection was then assessed for the occurrence of microsatellite instability (MSI). MSI was detected in 63/151 cases (42%) (Supplementary Data 1). It has been previously noted that a higher than expected fraction of CRC cell lines are MSI,

conceivably because MSI tumors can be more easily propagated *in vitro* ¹⁵. We also observed that the rare hypermutator/MSI-negative subgroup, identified in about 2% of CRC cases, is represented by three lines, HT115, HCC2998 and HT55 ¹⁶.

The mutational status of *KRAS*, *NRAS*, *BRAF*, and *PIK3CA* was next assessed as these genes are implicated in the response of CRC to EGFR targeted therapies and are often ascertained in the clinic (Supplementary Data 1). *KRAS*, *NRAS* and *BRAF* mutations were mutually exclusive and were present in 47%, 0.7% and 18% of the lines, respectively. *PIK3CA* mutations and PTEN status were detected in 18% and 14% of the cells, respectively, and occurred together with *KRAS* and *BRAF* at rates comparable to those previously reported^{17, 18, 19} (Supplementary Data 1). Overall, the complexity reached in this compendium effectively reflects the genetic heterogeneity recognized in sporadic CRCs.

Sensitivity to cetuximab is an intrinsic trait of CRC cells

The anti-EGFR monoclonal antibodies cetuximab or panitumumab are approved to treat metastatic CRC (mCRC) but achieve less than 10% objective response rates in unselected CRC patients when given in monotherapy ²⁰. Clinical benefit to EGFR blockade is confined to 25% of cases carrying tumors wild-type for *KRAS*, *NRAS* and *BRAF*²¹. To assess whether CRC lines recapitulated the above pharmacogenetic relationships, we determined sensitivity to cetuximab in the entire cell collection over a wide range of drug concentrations (Fig. 2, Supplementary Fig. 2 and Supplementary Data 1). Ten *RAS/BRAF* wild-type cell lines (7% of the entire collection or 20% of *RAS/BRAF* wild-type cells) were highly susceptible to cetuximab inhibition (Fig. 2, Supplementary Fig. 2a). Pharmacological parameters such as the area under the curve (AUC) or the percentage of cell growth inhibition at a clinically relevant drug concentration of 10 μ g/ml²² closely paralleled the

response rates observed in patients ^{20, 21}(Supplementary Fig. 3a). Cells classified as sensitive died upon anti-EGFR treatment (Supplementary Fig. 3 b,c,d), while partially sensitive cells were mainly growth impaired (Supplementary Fig. 3 e). As observed in the clinical setting, *KRAS, BRAF* or *NRAS* mutations conferred complete resistance to EGFR blockade in CRC cells (Fisher exact text p value<0.001) (Fig. 2).

CRC molecular subtypes are maintained in cell lines

Colorectal tumors can be classified in up to six unique subtypes with diverse clinical features according to the genes they express ^{8, 9, 10, 11, 12}. To what extent the transcriptional subtypes are cell-intrinsic or depend upon the microenvironment is unclear. Expression profiles were therefore used to assign each cell line to a molecular subtype ^{8, 9, 10, 11, 12} by the Nearest Template Prediction (NTP) algorithm, which also estimates the classification false discovery rate (FDR) ²³. Each of the five classifiers was able to assign (FDR < 0.2) the large majority of the cell lines to a subtype. Classifications of cell lines according to the five systems, together with the respective FDRs, are reported in Supplementary Data 2. Significant overlaps were detected among distinct classifiers. We found that the CRC-assigner (CRCA) classification system established by Sadanandam and colleagues¹² shared the highest overlap with all other classifiers (Supplementary Fig. 4a). It was therefore considered as a reference for reconciling the five classifiers into a unified consensus (Supplementary Fig. 4b).

Molecular and pharmacological associations, which previously emerged in CRC samples using the CRCA-based classification, were recapitulated in the cells collection (Fig. 3). MSI cells were significantly enriched in inflammatory and goblet subtypes (p-values: 0.0334 for inflammatory, 0.008 for goblet) and less prevalent in the Transit Amplifying (TA) and Stem

groups, which were instead mainly composed of microsatellite stable (MSS) cell lines (p-value<0.0005). *BRAF* mutated cell lines clustered in the inflammatory subtype (p-value = 0.0005), while *RAS* mutations were equally distributed among subgroups. Notably, 6 out of 9 *RAS/BRAF* wild type cetuximab sensitive cells belonged to the TA subtype (p=0.038). Such enrichment is in line with data on CRC specimens ¹² (Fig. 3).

Outlier tyrosine kinases overexpressed in CRC cells To unveil actionable targets likely to be driver oncogenes in CRC, we carried out an outlier expression analysis focused on protein kinases ²⁴, as the latter are frequently implicated in cancer and are ideally suited for therapeutic inhibition. Of 448 kinase genes expressed in the dataset ²⁵, 32 (7.1%) displayed an outlier expression profile. A striking increase in the prevalence of outliers was observed in the tyrosine kinase (TK) subfamily (15 out of 74, 2.84x enrichment, hypergeometric p-value<0.0001). Of these 15 TKs, 9 were outliers only in *RAS/BRAF* wild-type cells, and were therefore named «WT-specific». The enrichment for WT-specific outliers in the TK subgroup was significant (9 out of 15, 3.21x enrichment, hypergeometric p-value<0.0005). Interestingly, 8 out of the 9 WT-specific outlier TKs clustered in cell lines resistant to cetuximab. The list of WT-specific outliers' genes comprises *ALK, FER, FGFR2, KIT, NTRK1, NTRK2, PDGFRA* and *RET* (Fig. 4a).

Outlier kinase genes are therapeutic targets in CRC cells

We next performed genetic and functional validation of the TK outliers. We decided to focus on TKs for which targeted agents are in clinical trial or are already approved for treatment, namely ALK, FGFR2, KIT, NTRK1, NTRK2, RET and PDGFRA. This analysis revealed that overexpression of NTRK1 and FGFR2 is associated to molecular alterations, such as gene translocation (*NTRK1*) or gene amplification (*FGFR2*) (Supplementary Figs

5,6), previously described in cellular models and in cancer patients ^{26, 27, 28}, thus validating the experimental approach. The other tyrosine kinase outlier genes were not previously reported in CRC cells. We detected an in-frame gene fusion event between exon 20 of ALK and exon 13 of EML4 in C10 cells (Fig.4b). Break-apart FISH, cDNA PCR and western blot analysis confirmed the EML4-ALK translocation in C10 cells (Fig.4c-e). We reasoned that if the outlier kinase genes were causally implicated as oncogenic drivers, they should be functionally relevant in the corresponding cell models. We also postulated that the functional dependencies would be cell specific. To formally test these possibilities we used two complementary approaches: reverse genetics and pharmacological inhibition. Candidate-specific gene suppression was achieved with siRNA. In all cases reduced protein expression of the 'outlier' TK resulted in significant impairment of cell growth, which was often accompanied by downstream signaling inhibition and apoptosis (Fig. 4f; Supplementary Figs. 5-8).

Pharmacological validation of the TK outliers was performed using the following kinasetargeted drugs: CEP701 (TRK inhibitor), AZD4547 (FGFR2 inhibitor), crizotinib (ALK inhibitor), ponatinib (RET inhibitor) and imatinib and nilotinib (KIT and PDGFRA inhibitor). The drug inhibition profiles were cell specific and paralleled the expression profiles of individual TKs confirmed by RNAseq in the corresponding cell models (Fig. 4g-h, 5 and Supplementary Fig. 5-8). In two cases (KIT and PDGFRA) drug inhibition did not affect cell growth (Supplementary Fig. 8d,h). Nonetheless, in KIT overexpressing cells a reverse genetic experiment revealed functional dependency (Supplementary Fig. 8b,c).

Translocation of ALK and NTRK1 genes in colorectal cancer samples

To establish the clinical relevance of the cell line-based findings, we assessed whether the outlier TKs might be identified in CRC specimens using the same methodology. To this end we downloaded from cBbioPortal^{29, 30} (http://www.cbioportal.org) the RNAseq expression Z-scores for 352 CRC tissue samples generated by the Cancer Genome Atlas (TCGA) network ⁵. The entire set of candidates described above, with the exception of KIT, was independently validated as RNA expression outliers in the TCGA CRC dataset (Fig. 6a). As a further validation, we verified whether the 'outlier' strategy could lead to retrieval of TK candidates in archival CRC specimens for which RNA is typically not readily available. In this scenario, the most practical screening method should be capable of interrogating protein (rather than RNA) expression, on formalin-fixed paraffin embedded (FFPE) samples. As a proof of principle, we focused on NTRK1/Trk-A and ALK. An immunohistochemistry (IHC) based screen was applied to 742 CRC FFPE samples using a Trk-A (the protein encoded by NTRK1 gene) specific antibody. One Trk-A outlier sample was unequivocally identified (Fig. 6b,c). The IHC profile was followed by break-apart FISH analysis, which detected NTRK1 genetic rearrangement in the TrK-A positive sample (Fig. 6d). Screening with and ALK specific antibody of 742 CRC FFPE retrieved a sample with high levels of the ALK protein (Fig.6e,f). In the same specimen, break-apart FISH highlighted a genetic translocation involving the EML4 gene (Fig. 6g).

Discussion

Human neoplasms are highly heterogeneous, and clinical evidence indicates that pharmacogenetic relationships often involve oncogenic events occurring at low prevalence. Accordingly, precision oncology depends on the ability to functionally interrogate preclinical models capturing the molecular heterogeneity observed in patients. While collections of cancer cells derived from multiple tumor types have been previously described ^{1, 2, 4}, comprehensive databases of cells derived from single tumor lineages are much less common ^{16, 31}. To capture the clinical heterogeneity of CRCs, we assembled a collection of 151 cell lines. Notably, genetic finger printing and transcriptional profiling highlighted that several CRC lines previously thought to be unrelated are in fact derived from the same individual. This finding should be helpful in designing future studies and may facilitate the interpretation of previous analysis involving 'redundant' cell models.

The clinical relevance of the cell database was confirmed by showing that common oncogenic events (*KRAS, NRAS, BRAF, PIK3CA* and *PTEN* alterations) are present in the collection at rates similar to those found in CRC patients. Previous studies evaluated sensitivity to EGFR blockade (with the anti EGFR antibody cetuximab) in a fraction of the CRC lines described in this work ³². We found a strong concordance among measurements obtained in independent laboratories, indicating that sensitivity to EGFR inhibition is a stable phenotype of colorectal tumors. While the molecular determinants of EGFR sensitivity are presently unknown, our data suggest a prominent intrinsic (cell autonomous) component in colorectal tumors. We conclude that genetic and pharmacological profiling CRC lines can successfully nominate clinically relevant molecular determinants of response to targeted therapies.

In addition to the mutation-based categories described above, CRC can be subdivided in up to 6 transcriptional subtypes with distinct molecular and clinical features ^{8, 9, 10, 11, 12}, which we found to be robustly maintained in cells grown *in vitro*. This observation, along with the evidence that independent cell lines derived from the same patient consistently co-cluster transcriptionally, suggests that the transcriptional profile is stable and predominantly controlled by the genome in the absence of environmental inputs. These results further support the functional correspondence of CRC cell panel described here

with CRC specimens. Previous studies have also suggested that the transcriptional CRC subtypes may be associated with drug response. Cell models recapitulated the association between response to EGFR inhibitors and enrichment for the TA transcriptional class, which was identified in CRC clinical samples ¹². Along this line, the classification of cells reported in our work will facilitate future screenings with chemical or RNA interference libraries aimed at identifying novel pharmacogenomic relationships in specific transcriptional subtypes.

We previously reported that a subset of colorectal tumors, in which resistance to EGFR blockade occurs in a *RAS/RAF* wild type background, display overexpression of two tyrosine kinase genes, HER2 or MET ^{33, 34}. In 41/151 lines, which were analyzed in this study, the mechanism of resistance to cetuximab is unaccounted for. We hypothesized that in some of these cells resistance might likewise be driven by overexpression of a kinase gene. We therefore employed an outlier analysis based on expression levels of TK genes to identify additional kinase (and therefore actionable) targets in the subset of WT *RAS/BRAF* CRC cells intrinsically resistant to EGFR therapies.

Seven of the TKs retrieved by this approach (ALK, FGFR2, KIT, NTRK1, NTRK2, RET and PDGFRA) are the target of drugs undergoing clinical testing or approved. In some instances (FGFR2, ALK or NTRK1), overexpression of the TK gene was associated with amplification or translocation of the corresponding locus. The same genetic alterations have been previously reported in CRC cells and patients ^{26, 27, 28}, thus validating our approach. Our data further suggest that overexpression of TK sustains primary resistance to EGFR blockade, and could be used to identify patients unlikely to respond to cetuximab or panitumumab. Of note, with the exception of RET and FGFR2, which coexisted in NCIH716 cells, outlier expression of TKs was mutually exclusive.

Our results have broad implications for the use of kinase inhibitors in CRCs and potentially other cancer types. First, companion diagnostic assays are essential to identify those (rare) patients with tumors bearing aberrantly expressed kinases. Since the molecular events leading to protein overexpression can be multiple (gene amplification, gene fusions and translocations with different partners and breakpoints), we propose that IHC should be better suited than other methods to serve this purpose. As a test case, we demonstrated the feasibility of this approach to identify CRC patients carrying overexpressed and genetically altered Trk-A and ALK kinases. Additional work is warranted to develop and validate IHC assays for the other outlier TK genes. Importantly, given the low prevalence of these outlier events, strategies aimed at concomitantly screen for multiple molecular alterations should be implemented. Eventually, it is likely that next-generation sequencing will become the most effective approach to detect molecularly deregulated kinase genes as it allows concomitant assessment of single nucleotide variants as well as a gene copy number alterations and translocations.

Our results also highlight the relevance of clinical trials involving CRC patients carrying molecularly altered tyrosine kinase genes, which can be readily identified using FFPE specimens. Our data suggest that these patients are unlikely to benefit from anti EGFR antibodies but likely to profit from currently available drugs such as ALK and TRK inhibitors. It is important to note that in some instances (i.e. KIT), blockade of the kinase activity with a specific kinase inhibitor did not impair growth, while TK suppression with siRNA revealed functional dependency and triggered apoptosis. This suggests that down regulation of the candidate protein, for example through specific antibodies, might be necessary in some instances. Our data further suggest that tyrosine kinase inhibitors effectively impair tumor growth in cell lines in which transcriptional deregulation is

accompanied by genetic activation of the kinase (TRK, ALK, FGFR2), but not in the cell lines where no genetic activation was identified (PDGFR, KIT).

In conclusion our results describe a powerful preclinical resource capturing the heterogeneity of colorectal cancer patients, which can be used to define multidimensional information of clinical relevance and applicability.

METHODS

Cell lines and reagents

A collection of 152 cell lines of intestinal origin was assembled from different worldwide cell line banks or academic laboratories as indicated in Supplementary Data 1. All cell lines were maintained in their original culturing conditions according with supplier guidelines. Cells were ordinarily supplemented with FBS at different concentrations, 2mM L-glutamine, antibiotics (100U/mL penicillin and 100 mg/mL streptomycin) and grown in a 37° C and 5% CO₂ air incubator. Cells were routinely screened for absence of Mycoplasma contamination using the Venor[®] GeM Classic kit (Minerva biolabs). The identity of each cell line was checked by Cell ID[™] System and by Gene Print[®] 10 System (Promega), throught Short Tandem Repeats (STR) at 10 different loci (D5S818, D13S317, D7S820, D16S539, D21S11, vWA, TH01, TPOX, CSF1PO and amelogenin). Amplicons from multiplex PCRs were separated by capillary electrophoresis (3730 DNA Analyzer, Applied Biosystems) and analyzed using GeneMapperID software from Life Technologies. Resulting cell line STR profiles were cross-compared and matched with the available STR from ATCC, DSMZ, JCRB, Korean Cell Line Bank, ECCAC, and CellBank

Australia repositories online databases. Results of the analyzed loci for each cell line are provided in Supplementary Data 1.

Cetuximab was obtained from the Pharmacy at Niguarda Ca' Granda Hospital in Milan, Italy; Crizotinib, AZ4547, imatinib, ponatinib and nilotinib were purchased from Selleck Chemicals. CEP701 was purchased from Sigma Aldrich.

DNA analysis

Genomic DNA samples were extracted by Wizard® SV Genomic DNA Purification System (Promega). The MSI status has been evaluated by mean of MSI Analysis System kit (Promega). The analysis requires a multiplex amplification of seven markers including five mononucleotide repeat markers (BAT-25, BAT-26, NR-21, NR-24 and MONO-27) and two pentanucleotide repeat markers (Penta C and Penta D). The products were analyzed by capillary electrophoresis in a single injection (3730 DNA Analyzer, ABI capillary electrophoresis system (Applied Biosystems). The results were analyzed using GeneMarker V2.2.0 software. As cell lines do not have a corresponding normal tissue available for comparison, their profile has been analyzed with an MSS cell line of control, K562. Samples with instability in one, two or more markers (mononucleotide repeat) are defined as MS-instable (MSI). Samples with no detectable alterations are MS-stable (MSS). The same samples of genomic DNAs were used for PCR amplification to check 16 hot spot CRC mutations in all cells (reported in Supplementary Data 1). The mutational status of KRAS (exons 2, 3, 4), NRAS (exons 2, 3), HRAS (exons 2 and 3), BRAF (exon 15) and PIK3CA (exons 9 and 20) was determined by Sanger sequencing (primers' sequences are listed in Supplementary Data 3).

RNA extraction and gene expression profiling

RNA was extracted using miRNeasy Mini Kit (Qiagen), according to the manufacturer's protocol. The quantification and quality analysis of RNA was performed on a Bioanalyzer 2100 (Agilent), using RNA 6000 nano Kit (Agilent). Synthesis of cDNA and biotinylated cRNA was performed using the IlluminaTotalPrep RNA Amplification Kit (Ambion), according to the manufacturer's protocol using 500ng of total RNA. Quality assessment and quantification of cRNAs were performed with Agilent RNA kits on Bioanalyzer 2100. Hybridization of cRNAs (750 ng) was carried out using Illumina Human 48k gene chips (Human HT-12 V4 BeadChip). Array washing was performed using Illumina High Temp Wash Buffer for 10' at 55°C, followed by staining using streptavidin-Cy3 dyes (Amersham Biosciences). Probe intensity data were obtained using the Illumina Genome Studio software (Genome Studio V2011.1). Raw data were Loess normalized with the Lumi R package and further processed with Excel software..

Tissue of origin analysis

Thirty panels of genes with tissue-specific expression were retrieved from the TIGER portal¹³ (http://bioinfo.wilmer.jhu.edu/tiger/) and filtered to remove gene symbol redundancies. Similarly, to avoid probe redundancy in array data, when multiple probes matched the same gene we selected the one with the highest standard deviation. For the tissue of origin analysis, we assumed that the vast majority of the cell lines present in this dataset were of colorectal origin, while only a very small fraction could possibly derive from other tissues. Therefore, to increase sensitivity of the colon-specific gene panel, genes not expressed (L2S<9) in at least 50 of the 152 CRC cells were removed from the panel. To increase specificity of other tissue-specific gene panels, genes from these panels expressed (L2S>9) in more than three CRC cell lines (2%) were also removed. For each gene of a tissue-specific panel, the TIGER database provided an "EST Enrichment" (EE) score, proportional to the extent of tissue specificity. To give more weight to genes with

high EE score in the estimation of the tissue of origin, EE scores were squared and summed. Then, the squared EE score of each gene was divided by the sum, to obtain a "Scaled EE score" representing what fraction of the tissue specificity of the gene panel is brought by each individual gene. Finally, a "tissue score" was calculated for each gene panel by summing the products of gene expression and scaled EE score, and using the resulting value as exponential of 2 to obtain linear values.

Unsupervised expression clustering

After removing the COLO741 melanoma cell line from the dataset, we assessed global transcriptional correlations across the remaining 151 cell lines by unsupervised clustering A detection filter was applied to the 151 cell line dataset by requesting that the reported detection p-value reported in GSE59857 was zero in at least one of the 151 samples. After detection filtering, when multiple probes matched the same gene we selected the one with the highest standard deviation. This led to identification of 19,828 unique genes detected in at least one cell line. Before clustering,, we applied a further filter based on standard deviation (SD) across samples (SD of Log₂ Signal > 0.8). We then employed the Log₂ signal of 3,132 genes passing this filter for hierarchical clustering with maximum linkage based on Pearson correlation, using the GEDAS software ³⁵.

Drug proliferation assay

CRC cell lines were seeded at different densities $(2-4 \times 10^3 \text{ cells/well})$ in 100µl complete growth medium in 96-well plastic culture plates at day 0. The following day, serial dilutions of the indicated drugs were added to the cells in serum-free medium, while medium-only (in case of cetuximab) or DMSO-only (for all the other drugs) treated cells were included as controls. Plates were incubated at 37°C in 5% CO₂ for 4 or 5 days, after which cell viability was assessed by measuring ATP content through Cell Titer-Glo® Luminescent Cell Viability assay (Promega). Luminescence was measured by Perkin Elmer Victor X4. Based on cell viability after 4 days of treatment with clinical relevant concentration of 10µg/ml of cetuximab and on the shape of the drug screening-cell response growth curves, CRC cells were defined sensitive when cell viability was \leq 40% at 10µg/ml of cetuximab. We then calculated the area under the curve index (AUC) for each cell line of the panel and selected the threshold of AUC index \geq 13000 based on the values of AUC obtained in the subgroup of sensitive cell line. The area under the concentration—inhibition curve (Index AUC) was computed using Excel ³⁶. The "cetuximab effect arbitrary index" was calculated by subtracting the threshold value of 13000 from the AUC index of each cell line to highlight the subpopulation of sensitive cell lines (those with a negative index).

Classification of CRC cell lines into intrinsic subtypes

To classify cell lines into CRC subtypes according to previously published transcriptional classification systems, gene lists for the five classifiers were obtained from the supplementary information of the relevant publications. We then employed the NTP algorithm ²³ on the 151 CRC lines expression dataset of 19,828 genes described above. NTP is a nearest neighbor classifier developed to classify samples with defined gene signatures also when generated in a different microarray platform, and to provide an estimate of classification robustness by computing false discovery rate and p-value. NTP analysis was conducted using the GenePattern Bioportal³⁷ (http://www.broadinstitute.org/cancer/software/genepattern/).

For NTP implementation, we selected genes positively and specifically associated to one subtype. The CRC subtype classifiers were generated by different procedures. To perform reciprocal comparisons in an independent dataset, based on a different microarray platform, we had to adapt them to a common classification strategy. The thresholds for gene selection described below were chosen to remove genes with low class

discrimination power (to improve sensitivity and specificity) while at the same time maintaining a reasonable number of classifying genes (for the stability and platform independence of the classifier).

De Sousa E Melo⁹ and Sadanandam¹² gene signatures have originally been defined using the same approach, i.e. Significance Analysis of Microarrays³⁸followed by Prediction analysis for microarrays (PAM)³⁹, to build lists of genes positively or negatively associated to each molecular subtype. The extent and sign of association is represented by the PAM score. To select only genes with a strong positive association to one class, we associated each gene to the class corresponding to the maximum PAM value, and only if the difference between the maximum and second highest value was sufficiently high (greater than 0.1).

Subtype signatures of the Marisa¹⁰ classifier were defined by filtering the list of 1108 genes provided in Table S10 of their work, using the associated log fold changes and adjusted p-values for each gene in each of the six CRC subtypes. In particular, we assigned a gene to a class signature if two conditions were met: (i) its fold change in the class was the highest among its fold changes in all the classes; (ii) the adjusted p-value in the same class was the minimum between the adjusted p-values of all the other classes with a positive fold change. As done for De Sousa E Melo and Sadanandam classifiers, in order to select only genes with a strong association to a single subtype, we selected only those genes having a difference between the highest and second highest log fold change greater than 0.2. The Roepman ¹¹ signatures were defined on the basis of positive genes provided in Table 2 of the original work. The Budinska ⁸ gene signatures were provided on Table 1 of the original work as metagenes, each with statistical assessment of differential expression across subtypes. Adaptation to NTP required two steps: (i) selection of the metagenes specifically upregulated in each subtype, and (ii) extraction of the lists of genes belonging to the selected metagenes. A metagene was associated with a subtype if it was

significantly upregulated in that subtype, had the highest positive fold change (Log2 ratio vs median) in that subtype and if the difference against the second highest subtype was greater than a log2 ratio of 0.2. Then, for each subtype, the lists of genes belonging to the specifically upregulated metagenes were merged to build the subtype signatures. To compare the results obtained by the five different classifiers employed, we computed hypergeometric tests performing a paired-sample enrichment analysis, with Bonferroni correction of the resulting p-values. Associations between CRC subtypes and MSI status, *BRAF/KRAS* mutations and cetuximab sensitivity were evaluated by hypergeometric distribution analysis and Fisher's exact test.

Outlier expression analysis

Outlier analysis was performed starting from the 151 CRC lines expression dataset. Expression of a gene in a cell line was considered outlier if it was both >5 standard deviations and >5-fold greater than its median expression across all cells. Kinase-encoding genes were obtained from (<u>http://kinase.com/human/kinome/</u>)²⁵, after annotation with updated gene symbols and mapped on the detection-filtered expression dataset of 19,828 unique genes described above.

Gene copy number analysis

CRC cell lines were trypsinized, washed with PBS and centrifuged; pellets were lysed and DNA was extracted as described above. Real time PCR was performed with 10 ng of DNA per single reaction using GoTaq QPCR Master Mix (Promega) with an ABI PRISM[®] 7900HT apparatus (Applied Biosytems) (primers' sequences are available on request). Gene copy numbers were normalized to a control diploid cell line, HCEC ⁴⁰.

siRNA screening

The siRNA targeting reagents were purchased from Dharmacon, as a SMARTpool of four distinct siRNA species targeting different sequences of the target transcript. Cell lines were grown and transfected with SMARTpool siRNAs using RNAiMAX (Invitrogen) transfection reagents following manufacturer's instructions. Briefly RNAi screening conditions were as follow: on day one siRNA were distributed in each well of a 96-wellplate at final concentration of 20 nmol/L. Transfection reagents were diluted in OptiMEM and aliquoted at 10µl/well; after 20 minutes of incubation, 70µl of cells in media without antibiotics were added to each well. After five days, cell viability was estimated with a luminescent assay measuring cellular ATP levels with CellTiter-Glo[®] Luminescent Assay (Promega). Each plate included the following controls: mock control (transfection lipid only), AllStars (Qiagen) as negative control; Polo-like Kinase 1 (PLK1) (Dharmacon) as positive control⁴¹.

RNA seq

Total RNA extracted from C10, KM12, SNU503, HCA24, CACO2, and, NCIH716 cells was processed for RNAseq analysis with Ion TotRNA Seq kit v2[®] following manifacturer's instruction and sequenced with the Ion Proton[™] System (Ambion - Life Technology). Each fastq file was realigned using Tophat2 v2.0.11⁴². Version hg19 of the genome was used and Gencode v19 as reference transcriptome database. The average value of the reads aligned by Tophat2 is 61%. The Fragments Per Kilobase of exon per Million fragments mapped (FPKM) was computed on the aligned reads using Cufflinks software v2.2.1 ⁴³. Moreover, each fastq file was analyzed by FusionMap software v 7.0.1.25 ⁴⁴ obtaining the list of fusion transcripts for each cell line.

Immunohistochemistry.

FFPE colorectal cancer samples were available through the clinical database of Ospedale Niguarda and Istituto Nazionale Tumori Regina Elena. Additional CRC specimens tissue microarrays were purchased from Abcam (Cat. ab128127). Trk-A and ALK protein expression was evaluated by immunohistochemistry (IHC) performed on 3-4µm thick tissue sections, using specific antibodies In order to assess Trk-A expression, the monoclonal antibody (Clone ID EP1058Y, Epitomics, dilution 1:200), which recognizes the C-terminal portion of the protein was used and the analysis was performed with the automated system BenchMark Ultra (Ventana Medical System, Inc., Roche) according to the manufacturer's instructions, with minimum modifications. Samples were incubated with the primary antibody for 1h at 42°C, after which a chromogenic reaction was performed using the ultraView Universal Alkaline Phosphatase Red Detection kit (Ventana, Roche). ALK protein espression was detected using the anti ALK mouse monoclonal antibody (clone 5A4) which recognizes the intracellular region of ALK. The analysis was performed manually, using N-Histofine ALK Detection Kit, according to the manufacturer's nstructions (Nichirei Biosciences Inc). Samples were incubated with the primary antibody for 30 min at room temperature and DAB was used as chromogen. For both proteins, specimens were considered positive when the expression was strong and cytoplasmic and present in at least 10 % of cells. Healthy tissue, i.e., normal colon mucosa, was used as internal negative control. Images were captured with the AxiovisionLe software (Zeiss) using an Axio Zeiss Imager 2 microscope (Zeiss).

Fluorescent in situ hybridisation analysis.

Tissue sections and cyto-embedded cell lines for Fluorecent in situ hybridization (FISH) experiment were prepared according to the manufacturer's instructions of Histology FISH Accessory kit (Dako). For both types of samples (tissue and cell lines) the last steps before

hybridization were: dehydration in ethanol series (70%, 90%, 100%), 3 washes (5' each) and air drying. To identify possible NTRK1 gene rearrangements, a dual colour break apart probe has been designed a contiguous of two BAC (Bacterial Artificial Chromosome) (http://genome.ucsc.) probes, RP11-349I17(1g23.1) and RP11-1038N13 (1g23.1), all together spanning an approximately 335 kb region encompassing the NTRK1 gene. A dual colour FISH analysis was performed using respectively: 10 µl mix- probe made up by 1.5µl BAC genomic probe RP11-349I17 (1g23.1) labelled in Spectrum Orange (Empire Genomics), 1,5 µl BAC genomic probe RP11-1038N13 (1g23.1) labelled in SpectrumGreen, and 7 µI LSI-WCP hybridisation buffer (Vysis) for each slide. In order to identify possible ALK gene rearrangements, an LSI ALK (2p23) Dual Color, Break Apart (Vysis) probe was used. Probes and target DNA of specimens were co-denatured in HYBRite System (DakoGlostrup) respectively at 75°C for 5 min using NTRK1 probe and at 73° for 5 min for ALK probe, therefore hybridized overnight at 37° C. Slides were washed with post-hybridisation buffer (DakoGlostrup) at 73°C for 5 min and counterstained wit 4,6diamino-2phenylindole (DAPI II; Vysis, Downers Grove, IL. USA). Fluorescent in situ hybridisation signals were evaluated with a Zeiss Axioscope Imager. Z1 (Zeiss) equipped with single and triple band pass filters. Images for documentation were captured with CCD camera and processed using the MetaSystems Isis software. Samples were scored as positive either for NTRK1 or ALK gene's rearrangement when a split signal of the probes was observed, in at least 15% of 100 cells analyzed in 10 different fields. Healthy tissue, i.e, normal colon mucosa, was used as internal negative control.

Western blotting analysis

Prior to biochemical analysis, all cells were grown in their specific media supplemented with 10% FBS. Total cellular proteins were extracted by solubilizing the cells in EB buffer (50 mMHepes pH 7.4, 150 mMNaCl, 1% Triton X-100, 10%

glycerol, 5 mM EDTA, 2 mM EGTA; all reagents were from Sigma-Aldrich, except for Triton X-100 from Fluka) in the presence of 1 mM sodium orthovanadate, 100 mM sodium fluoride and a mixture of protease inhibitors. Extracts were clarified by centrifugation, normalized with the BCA Protein Assay Reagent kit (Thermo). Western blot detection was performed with enhanced chemiluminescence system (GE Healthcare) and peroxidase conjugated secondary antibodies (Amersham). The following primary antibodies were used for western blotting (all from Cell Signaling Technology, except where indicated): anti-FGFR2 (1:1000) ; anti-ALK; anti-Trk-A (Santa Cruz) (1:500); anti-Trk (Santa Cruz) (1:500); anti-cKIT (1:1000); anti-PDGFRA (1:1000); anti-phospho-p44/42 ERK (Thr202/Tyr204) (1:1000); anti-p44/42 ERK (1:1000); anti-phospho-MEK1/2 (Ser217/221) (1:1000); anti-MEK1/2 (1:1000); antiphospho AKT (Ser473) (1:1000); anti-AKT (1:1000); anti-PARP (1:1000); anti-actin (Santa Cruz) (1:3000); anti-vinculin (Millipore) (1:5000); anti-PTEN (1:1000). Anti-RET antibody⁴⁵ (1:500) was kindly provided by Prof. Massimo Santoro (Federico II, Naples). The full-length western blot membranes are shown in Supplementary Figs 9–13.

References

- 1. Barretina J, *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607 (2012).
- 2. Garnett MJ, *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570-575 (2012).
- 3. Park ES, *et al.* Integrative analysis of proteomic signatures, mutations, and drug responsiveness in the NCI 60 cancer cell line set. *Mol Cancer Ther* **9**, 257-267 (2010).
- 4. Basu A, *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* **154**, 1151-1161 (2013).
- 5. Network CGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337 (2012).

- 6. Smith CG, *et al.* Somatic profiling of the epidermal growth factor receptor pathway in tumors from patients with advanced colorectal cancer treated with chemotherapy ± cetuximab. *Clin Cancer Res* **19**, 4104-4113 (2013).
- 7. Roth AD, *et al.* Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *J Natl Cancer Inst* **104**, 1635-1646 (2012).
- 8. Budinska E, *et al.* Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J Pathol* **231**, 63-76 (2013).
- 9. De Sousa E. Melo F, *et al.* Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat Med* **19**, 614-618 (2013).
- 10. Marisa L, *et al.* Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* **10**, e1001453 (2013).
- 11. Roepman P, *et al.* Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *Int J Cancer* **134**, 552-562 (2014).
- 12. Sadanandam *A, et al.* A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med* **19**, 619-625 (2013).
- 13. Liu X, Yu X, Zack DJ, Zhu H, Qian J. TiGER: a database for tissue-specific gene expression and regulation. *BMC Bioinformatics* **9**, 271 (2008).
- 14. Vécsey-Semjén B, *et al.* Novel colon cancer cell lines leading to better understanding of the diversity of respective primary cancers. *Oncogene* **21**, 4646-4662 (2002).
- 15. Wilding JL, Bodmer WF. Cancer cell lines for drug discovery and development. *Cancer Res* **74**, 2377-2384 (2014).
- 16. Mouradov D, *et al.* Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Res*, (2014).
- 17. Barault L, *et al.* Mutations in the RAS-MAPK, PI(3)K (phosphatidylinositol-3-OH kinase) signaling network correlate with poor survival in a population-based series of colon cancers. *Int J Cancer* **122**, 2255-2259 (2008).
- 18. De Roock W, *et al.* Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* **11**, 753-762 (2010).
- 19. Sartore-Bianchi A, *et al.* Multi-determinants analysis of molecular alterations for predicting clinical benefit to EGFR-targeted monoclonal antibodies in colorectal cancer. *PLoS ONE* **4**, e7287 (2009).
- 20. Cunningham D, *et al.* Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N Engl J Med* **351**, 337-345 (2004).

- 21. De Roock W, De Vriendt V, Normanno N, Ciardiello F, Tejpar S. KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. *Lancet Oncol* **12**, 594-603 (2011).
- 22. Tabernero J, *et al.* Cetuximab administered once every second week to patients with metastatic colorectal cancer: a two-part pharmacokinetic/pharmacodynamic phase I dose-escalation study. *Ann Oncol* **21**, 1537-1545 (2010).
- 23. Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. *PLoS One* **5**, e15543 (2010).
- 24. Kothari V, *et al.* Outlier kinase expression by RNA sequencing as targets for precision therapy. *Cancer Discov* **3**, 280-293 (2013).
- 25. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934 (2002).
- 26. Lin E, *et al.* Exon array profiling detects EML4-ALK fusion in breast, colorectal, and non-small cell lung cancers. *Mol Cancer Res* **7**, 1466-1476 (2009).
- 27. Ardini E, *et al.* The TPM3-NTRK1 rearrangement is a recurring event in colorectal carcinoma and is associated with tumor sensitivity to TRKA kinase inhibition. *Mol Oncol*, (2014).
- 28. Mathur A, Ware C, Davis L, Gazdar A, Pan BS, Lutterbach B. FGFR2 Is Amplified in the NCI-H716 Colorectal Cancer Cell Line and Is Required for Growth and Survival. *PLoS One* **9**, e98515 (2014).
- 29. Gao J, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1 (2013).
- 30. Cerami E, *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404 (2012).
- 31. Beaufort CM, *et al.* Ovarian Cancer Cell Line Panel (OCCP): Clinical Importance of In Vitro Morphological Subtypes. *PLoS One* **9**, e103988 (2014).
- 32. Ashraf SQ, Nicholls AM, Wilding JL, Ntouroupi TG, Mortensen NJ, Bodmer WF. Direct and immune mediated antibody targeting of ERBB receptors in a colorectal cancer cell-line panel. *Proc Natl Acad Sci U S A* **109**, 21046-21051 (2012).
- 33. Bertotti A, *et al.* A molecularly annotated platform of patient-derived xenografts ("xenopatients") identifies HER2 as an effective therapeutic target in cetuximabresistant colorectal cancer. *Cancer Discov* **1**, 508-523 (2011).
- 34. Bardelli A, *et al.* Amplification of the MET receptor drives resistance to anti-EGFR therapies in colorectal cancer. *Cancer Discov* **3**, 658-673 (2013).

- 35. Fu L, Medico E. FLAME, a novel fuzzy clustering method for the analysis of DNA microarray data. *BMC Bioinformatics* **8**, 3 (2007).
- 36. Whitehouse PA, *et al.* Heterogeneity of chemosensitivity of colorectal adenocarcinoma determined by a modified ex vivo ATP-tumor chemosensitivity assay (ATP-TCA). *Anticancer Drugs* **14**, 369-375 (2003).
- 37. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet* **38**, 500-501 (2006).
- 38. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121 (2001).
- 39. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* **99**, 6567-6572 (2002).
- 40. Roig AI, *et al.* Immortalized epithelial cells derived from human colon biopsies express stem cell markers and differentiate in vitro. *Gastroenterology* **138**, 1012-1021.e1011-1015 (2010).
- 41. Brough R, *et al.* Functional viability profiles of breast cancer. *Cancer Discov* **1**, 260-273 (2011).
- 42. Trapnell C, *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-515 (2010).
- 43. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111 (2009).
- 44. Ge H, Liu K, Juan T, Fang F, Newman M, Hoeck W. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. *Bioinformatics* **27**, 1922-1928 (2011).
- 45. De Falco V, *et al.* Ponatinib (AP24534) is a novel potent inhibitor of oncogenic RET mutants associated with thyroid cancer. *J Clin Endocrinol Metab* **98**, E811-819 (2013).
- 46. Valtorta E, *et al.* KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. *Int J Cancer* **133**, 1259-1265 (2013).

Author Contributions

A.B., E.M., F.D.N., M.R. and G.P. designed the study. M.R., G.P., C.C., E.V., M.B. S.L.,

B.M.S.V., E.M. and F.D.N performed experiments and analyzed the data. G.C., C.I., M.B.,

M.M. and F.C., contributed to the analysis of data. M.L. established HROC panel of CRC

cells. A.B., F.D.N., E.M., M.R. and G.P. wrote the manuscript. A.B., F.D.N. and E.M. supervised the study.

Acknowledgments

We thank Lazzari L., Lallo A., Misale S., Arena S. and Rossi T. for helping in the cetuximab pharmacological screening; Cantarella D. and Porporato R. for microarray expression profiling. We thank Prof. Massimo Santoro for supplying RET antibody. We are grateful to the researchers named in Table S1 for the kind gift of the cell lines. Supported by the European Community's Seventh Framework Programme under grant agreement no. 602901, MErCuRIC (A.B.); Associazione Italiana per la Ricerca sul Cancro (AIRC) IG grant no. 12812 (A.B.) and no. 12944 (E.M.); AIRC MFAG no. 11349 (F.D.N.); grant "Farmacogenomica"—5 per mille 2009 MIUR—Fondazione Piemontese per la Ricerca sul Cancro—ONLUS (F.D.N.); AIRC 2010 Special Program Molecular Clinical Oncology 5 per mille, project no. 9970 (A.B. and E.M.); FPRC 5 per mille 2010 and 2011 Ministero della Salute (A.B., F.D.N. and E.M.); Ministero dell'Istruzione, dell'Università e della Ricerca, progetto PRIN (A.B.); Progetti di Ateneo-2011, Università di Torino (ORTO11RKTW to A.B.).

Accession codes

Gene expression microarray data have been deposited in the GEO database with accession number GSE59857

Competing financial interests

Authors declare NO competing financial interests

Figure Legends

Figure 1. Expression clustering of 151 CRC cell lines.

Unsupervised hierarchical clustering of 151 CRC cell lines based on their global expression profile. Cell lines derived from the same individual (same STR) are highlighted by the same color. Names of cell lines with identical STR are reported below the dendrogram in the order by which they occur in the cluster, from left to right.

Figure 2. Cetuximab screening on CRC cell lines. The indicated cell lines were treated with increasing concentrations of cetuximab for 4 days and cell viability was assessed by measuring ATP content. Bars represent an arbitrary index of cetuximab effect on each cell line as detailed in the methods. Cell lines sensitive to cetuximab treatment are shown with a negative index. Red bars represent *KRAS* altered lines; yellow bars indicate *NRAS* mutated cells; blue bars indicate genetic alterations affecting codon V600 of *BRAF*; black bars indicate *RAS/BRAF* wild type cells. NCIH630 cells are KRAS amplified ⁴⁶.

Figure 3. CRC transcriptional subtypes are maintained in cell lines. CRC cell lines are classified according to the "CRC-assigner" intrinsic transcriptional subtypes and the colon crypt location previously described¹². MSI status, *RAS/BRAF* mutations and cetuximab sensitivity are indicated. FDR: false discovery rate. The Fig. includes only CRC cell lines that were assigned to a subtype with FDR < 0.2. Among cell lines sharing the same STR profile we selected, the original cell line in case of groups of cells including also derived clones, and in the absence of such information, the one classified with the highest confidence by the NTP algorithm..RAS_M: *RAS* mutated cells. BRAF_M: *BRAF* mutated cells. Cetux: cetuximab.

Figure 4. Identification of outlier kinase genes in 151 CRC cell lines. (a) Scatter plot representation of transcriptional outlier kinases in 151 CRC cells. Colored circles represent outlier kinases with >5 standard deviations and >5-fold differential expression compared to median expression. Asterisk labels the KM12 family of cells: KM12, KKM12C, KM12SM, KM12L4. (b) RNA seq analysis unveiled an in-frame gene fusion event between exon 20 of the ALK gene and the exon13 of the EML4 gene on chromosome 2 in the C10 CRC cell line. (c) Break apart FISH analysis of EML4-ALK in C10 cells shows clear separation of green (5') and red (3') signals corresponding to the ALK gene (original 63x magnification scale bar=10µm). Scale bars, 10µm (d) PCR on the cDNA of C10 cells confirmed the EML4-ALK translocation (e) Western blot analysis confirmed overexpression of ALK protein in C10 cells, at a molecular weight compatible with an EML4-ALK genetic rearrangement. (f) RNAi knockdown of ALK inhibits cell proliferation of C10 cell line harboring the EML4-ALK fusion gene. C10 cells were analyzed by ATP lite proliferation assay 5 days after transfection with ALK-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). Data are expressed as average ± s.d. of three independent experiments. si-RNA mediated downregulation of ALK levels were verified by western blot analysis. (g) C10 cells, resistant to the EGFR inhibitor cetuximab, are sensitive to the ALK inhibitor crizotinib. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average \pm s.d. of three independent experiments. (h) Crizotinib treatment downregulates MAPK and PI3K pathways in C10 cells harboring the ALK genetic rearrangement. C10 cell lines were treated with ALK inhibitor 1µM crizotinib for 12h, after which whole-cell extracts were subjected to western blot analysis.

Figure 5. Overexpressed kinase genes are therapeutic targets in CRC. (a) RNA seq analysis of the indicated cell lines shows overexpression of TK outlier originally identified by microarray data analysis. FPKM: Fragments Per Kilobase of exon per Million fragments mapped. **(b)** Drug profiling with kinase inhibitors. The indicated cell lines were treated with crizotinib (ALK inhibitor) 240nM, CEP701 (TRK inhibitor) 100nM, AZD4547 (FGFR inhibitor) 100nM, ponatinib⁴⁵ (RET inhibitor) 100Nm or nilotinib (KIT and PDGFRA inhibitor) 240nM for 5 days. Viability was then assessed by measuring ATP content.

Fig. 6. NTRK1 and ALK molecular alterations in CRC specimen. (a) Scatter plot representation of transcriptional outlier kinases in 352 CRC tissue samples from the TCGA database. Red circles highlight with Z-scores samples values>6. (b-c) Immunohistochemistry based identification of FFPE CRC displaying Trk-A protein overexpression. 1/742 samples was markedly positive for Trk-A overexpression. In panel b original 4x magnification, scale bar=500µm; in panel **c** original 20x magnification, scale bar=100µm.(d) Break-apart FISH analysis revealed NTRK1 gene rearrangement in the FFPE sample positive by IHC (original 63x magnification, scale bar=10µm). (e-f) Immunohistochemistry based identification of FFPE CRC displaying ALK protein overexpression. 1/742 samples was markedly positive for ALK overexpression. In panel e original 4x magnification, scale bar=500 μ m; in panel f original 20x magnification, scale bar=100µm) (g) Break-apart FISH analysis revealed ALK gene rearrangement in the FFPE sample positive by IHC (original 63x magnification, scale bar=10µm).



COLO320DM	LS174T	GP5D	CX1	KM12SM	COGA5L	COLO201	SW620	НСТ8
COLO320HSR	LS180	GP2D	WIDR	KM12L4	COGA5	COLO205	SW480	HCT15
COLO320			HT29	KM12C				DLD1
			KM20	KM12				MIP101
								C146
								C170







Log [M] Drug

а













е



f





а





Supplementary Figure 1: Tissue of Origin analysis on 152 cell lines. (a) Heatmap representation of the 30 "Tissue scores" for the 152 cell lines. The scores summarize the global expression of the tissue specific genes for each one of the 30 tissues analyzed. (b) Plot representation of the weighed expression of the colon genes versus skin genes. Circle highlights COLO741 cells showing high "skin" score. (c) Log2 expression of the skin-specific tyrosinase (*TYR*) gene in 152 cell lines. Circle highlights COLO741 cells showing an outlier expression profile for the *TYR* gene. (d) Plot representation of the weighed expression of the colon genes versus small intestine genes. Circle highlights Hutu80 cells displaying high "small intestine" score.



Supplementary Figure 2. Cetuximab response curves in 150 CRC cell lines. 150 cell lines of colorectal origin were treated with increasing concentrations of cetuximab for 4 days and cell viability was assessed by measuring ATP content. Data are expressed as percent viability compared with medium-only treated cells. Based on their response cells can be classified in sensitive (red lines); partially sensitive (green lines) and resistant (blue lines) to cetuximab treatment.



Supplementary Figure 3. Cetuximab effect on CRC cell lines. (a) Cell viability of each CRC cell line at clinical relevant concentration of 10µg/ml of cetuximab after 4 days of treatment (indicated by yellow dots) closely parallel AUC values. **(b-e)** Analysis of cell viability after 4-5 days of cetuximab treatment indicates net loss of cells compared to the number of cells at time "0", in three representative 'sensitive' cell models, even at clinical relevant dose of 10 µg/ml of cetuximab **(b-d)**; while cetuximab exerts mainly a cytostatic effect (reduction of growth rate) on the intermediate cell lines LIM1215 **(e)**. Data are expressed as percent viability compared to the number of cells at day "0".

Sadanandam et al.

		Enterocyte	Inflammatory	Goblet	TA	Stem
	C6	3,06E-06	21,959	9,001	14,511	3,761
	C2	14,339	1,02E-04	5,338	3,875	6,971
	C3	17,441	1,925	2,16E-08	8,854	1,925
Marisa et al.	C1	16,782	18,249	12,326	1,67E-04	18,249
	C5	22,491	19,428	2,050	6,94E-06	3,073
	C4	3,226	14,779	1,727	0,046	9,08E-17
	Subtype E	7,296	33,981	42,886	32,523	17,982
	Subtype C	18,906	1,21E-09	0,993	2,663	1,643
Budinska et al.	Subtype A	4,817	16,797	1,47E-02	24,530	9,817
	Subtype B	16,474	0,499	16,267	8,21E-15	0,089
	Subtype D	18,379	8,072	3,959	0,682	1,72E-17
•						
	CCS2	20,814	6,17E-05	8,92E-03	6,272	7,253
De Sousa E Melo et al.	CCS1	0,135	10,647	21,772	6,48E-14	0,072
	CCS3	11,871	21,929	14,085	0,153	7,93E-18
-						
	A-Type	4,331	5,78E-04	2,73E-02	2,743	8,924
Roepman et al.	B-Type	3,869	10,139	18,358	2,99E-16	0,595
	C-Type	22,691	19,054	29,743	0,294	1,37E-13
-						

b



Supplementary Figure 4: Expression based classification of CRC cell lines. (a) CRCassigner (CRCA) is the reference classifier. Enrichment matrix (Bonferroni-corrected hypergeometric p-values) comparing members of CRCA subtype versus others expression-based CRC classifiers. Orange highlighting indicates enrichment p-values< 0.05, green highlighting indicates depletion p-values< 0.05. (b) Graphic representation of the consensus across the intrinsic subtypes identified in cell lines by CRCassigner (CRCA) by Sadanandam et al. and other four different classification systems. The number of cell lines significantly assigned to each subtype (FDR<0.2) is reported.

а



Supplementary Figure 5. Characterization of oncogenic *TPM3-NTRK1* gene fusion in KM12 cells. (a) RNA seq analysis identified gene fusion event between exon 7 of the *TPM3* gene and the exon10 of *NTRK1* gene in the KM12 CRC cell line. (b) PCR on the cDNA of KM12 cell lines confirmed the *TPM3-NTRK1* gene rearrangements. KM12SM cells, a derivate of the KM12 cells, were used as positive control. (c) Break apart FISH analysis of *TPM3-NTRK1* in KM12 cells shows clear separation of green (5') and red (3') signals corresponding to the *NTRK1* gene. Scale bars, 10µm. (d) Western blot analysis confirmed overexpression of the TRKA protein (codified by *NTRK1* gene) in KM12 cells, at a molecular weight compatible with *TPM3-NTRK1* translocation. (e) RNAi knockdown of NTRK1 inhibits cell proliferation of KM12 cells harboring the *TPM3-NTRK1* fusion gene. KM12 cells were analyzed by ATP lite proliferation assay 5 days after transfection with NTRK1-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). Data are expressed as average ± s.d. of three independent experiments. si-RNA mediated downregulation of NTRK1 levels were verified by western blot. (f) KM12 cells, resistant to the EGFR inhibitor cetuximab, are sensitive to the TRK inhibitor CEP701. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average ± s.d. of three independent experiments. KM12 cells were treated with 1µM CEP701 for 12h, after which whole-cell extracts were subjected to western blot analysis.



Supplementary Figure 6. Oncogenic addiction to *FGFR2* genetic amplification. (a) gDNA RTPCR identified *FGFR2* gene copy number increase in the NCIH716 CRC cell line. (b) Western blot analysis confirmed overexpression of FGFR2 protein in the NCIH716 cells. (c) RNAi knockdown of FGFR2 inhibits NCIH716 cell proliferation overexpressing the tyrosine kinase receptor. Cells were analyzed by ATP lite proliferation assay 5 days after transfection with FGFR2-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). Data are expressed as average ± s.d. of three independent experiments. si-RNA mediated downregulation of FGFR2 levels were verified by western blot analysis. (d) NCIH716 cell lines, resistant to EGFR inhibitor cetuximab, are sensitive to FGFR inhibitor AZD4547. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average ± s.d. of three independent experiments. (e) AZD4547 treatment downregulates MAPK and PI3K pathways in NCIH716 cells harboring *FGFR2* gene copy number increase. NCIH716 cells were treated with 0.5µM AZD4547 for 12h, after which whole-cell extracts were subjected to western blot analysis.



Supplementary Figure 7. Oncogenic addiction to overexpressed NTRK2 and RET kinases. (a) RNAi knockdown of NTRK2 inhibits cell proliferation of SNU503 cell line. SNU503 cells were analyzed by ATP lite 5 days after transfection with NTRK2-specific pooled siRNAs, or transfection reagent (mock). Data are expressed as average \pm s.d. of three independent experiments. si-RNA mediated downregulation of NTRK2 levels were verified by cDNA RTPCR. (b) SNU503 cells, resistant to the EGFR inhibitor cetuximab, are sensitive to the TRK inhibitor CEP701. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average \pm s.d. of three independent experiments. (c) Treatment with TRK inhibitor CEP701 initiates apoptosis in NTRK2-dependent SNU503 CRC cells. Protein lysates were analyzed by western blot after 3 days of treatment with 1µM CEP701. (d) Western blot analysis confirmed overexpression of RET protein in NCIH716 cells. (e) NCIH716 cells, resistant to the EGFR inhibitor cetuximab, are sensitive to the RET inhibitor ponatinib and to the FGFR inhibitor AZD4547. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average \pm s.d. of three independent experiments. (f) RNAi knockdown of RET inhibits cell proliferation in the NCIH716 cells. NCIH716 cells were analyzed by ATP lite proliferation assay 5 days after transfection with RET-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). Data are expressed as average \pm s.d. of three independent experiments. si-RNA mediated downregulation of RET levels were verified by western blot analysis.



Supplementary Figure 8. Functional validation of KIT and PDGFRA overexpression. (a) Western blot analysis confirmed overexpression of KIT protein in HCA24 cell lines. (b) RNAi knockdown of KIT inhibits cell proliferation in HCA24 cells. HCA24 cells were analyzed by ATP lite 5 days after transfection with KIT-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). Data are expressed as average ± s.d. of three independent experiments. si-RNA mediated downregulation of KIT levels were verified by western blot. (c) RNAi knockdown of KIT initiates apoptosis in KIT-dependent HCA24 CRC cell lines. Protein lysates were analyzed by western blot 3 days after transfection with KIT-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). (d) HCA24 cells, resistant to the EGFR inhibitor cetuximab, are resistant also to the KIT kinase inhibitors nilotinib or imatinib. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average ± s.d. of three independent experiments. (e) Western blot analysis confirmed overexpression of PDGFRA protein in CACO2 cell lines. (f-g) CACO2 cells are not addicted to the PDGFRA overexpression. RNAi knockdown of PDGFRA failed to induce apoptosis. Protein lysates were analyzed by western blot 3 days after transfection with PDGFRA-specific pooled siRNAs, scrambled siRNA, or transfection reagent (mock). (h) CACO2 cell lines, resistant to the EGFR inhibitor cetuximab, are resistant also to the PDGFRA kinase inhibitors nilotinib or imatinib. Cell viability was assessed by measuring ATP content after 5 days of treatment. (h) CACO2 cell lines, resistant to the EGFR inhibitor cetuximab, are resistant also to the PDGFRA kinase inhibitors nilotinib or imatinib. Cell viability was assessed by measuring ATP content after 5 days of treatment. Data are expressed as average ± s.d. of three independent experiments.



Supplementary Figure 9. Full lengths western blot of Figure 4. (4e) C10 CRC cells overexpress ALK protein. (4f) RNAi knockdown of ALK inhibits cell proliferation of C10 cell line harboring the EML4-ALK fusion gene. (4h) Crizotinib treatment downregulates MAPK and PI3K pathways in C10 cells.



S5g



Supplementary Figure 10. Full lengths western blot of Supplementary Figure 5. (S5d) KM12 CRC cells overexpress TRKA protein. (S5e) RNAi knockdown of TRKA inhibits cell proliferation of KM12 cell line harboring the TPM3-NTRK1 fusion gene. (S5g) CEP701 treatment downregulates MAPK and PI3K pathways in KM12 cells.



Supplementary Figure 11. Full lengths western blot of Supplementary Figure 6. (S6b) NCIH716 CRC cells overexpress FGFR2 protein. (S6c) RNAi knockdown of FGFR2 inhibits cell proliferation of NCIH716 cell line harboring the FGFR2 gene amplification. (S6e) AZD4547 treatment downregulates MAPK and PI3K pathways in NCIH716 cells.



Supplementary Figure 12. Full lengths western blot of Supplementary Figure 7. (S7c) Treatment with TRK inhibitor CEP701 initiates apoptosis in NTRK2-dependent SNU503 CRC cells. (S7d) NCIH716 CRC cells overexpress RET protein. (S7f) RNAi knockdown of RET inhibits cell proliferation of NCIH716 cell line harboring the RET overexpression



Supplementary Figure 13. Full lengths western blot of Supplementary Figure 8. (S8a) HCA24 CRC cells overexpress KIT protein. (S8b) RNAi knockdown of KIT inhibits cell proliferation of HCA24 cells. (S8c) RNAi knockdown of KIT initiates apoptosis in KIT-dependent HCA24 CRC cells. (S8e) CACO2 CRC cells overexpress PDGFRA protein. (S8bf) RNAi knockdown of PDGFRA does not affect cell proliferation of CACO2 cells. (S8g) RNAi knockdown of PDGFRA does not induce apoptosis in CACO2 CRC cells.

Supplmentary Dat and acquisition so a brief description	a 1. Molecular and pha arces, (ii) STR profiles, (of the section content. I	armacological annotat (III) MSI and mutationa Footnotes list contact d	ion of 151 CRC cell lines. I annotations and (iv) pha etails of scientists who pro	This table is sub: macological ann ilded the indicat	divided in four solutions. On t ad cell lines.	r wctions, separ top of the wctic	ing: () identification ns, a note provides de	ame alls and																								
				CRC cell line profiles were the same colo	is identity was cross-compan or.	s checked by Sh ed and matched	nt Tandem Repeats (5 with the available AT	TR) profiles throug CC STR, DSMZ, J	h simultaneousi 2RB, Korean Ce	y amplification of I Line Bank, ECC	ten STR loci (D AC, and CellBar	ssata, D135317, D7 k Australia repositori	5820, D165539, D scotline database	21511, WA, TH & Cell lines derive	1, TPOX, CSF1 d from the sam	PO and amelog a individual (sim	enin). Resulting cell lar STR) are highlig	line STR List o phied by anal anal	of the geneti ried hotspote yeas, is also i	tic status of the 151 CRC cell is. MSIMSS status are report indicated. Positive: PTEN p	I lines for the HRA ted. PTEN protei rotein detected. N	S, NRAS, KRA n expression, ev ull: PTEN prote	S, BRAF and PIK3CA valueted by weden blot ein absent.	CRC cells w each cell la percentage	ere treated with i the the area unde (%) of growth init	indicated concentration in the concentration-in hibition at indicated co	s of cetuximab f ibition curve (In centrations is lis	or 4 days, after tex AUC) (see I ted. Cetux: cet	that cell viabil Methoda) is rep usimab	ty was a memory sted. For each	d by ATP con I drug concen	tent. For tration the
CRC Cell Line Name	Supplier ID	Cat ID	Date of Arrival	Locus AMEL	Locu	a CSF1PO	Locus DSS8		ocum 075828	Locus	D135317	Locus D165539	Locus 0215	11 Locus	THD1	ocus TPOX	Locus vW	. 15	i / MSS H	AA change HRAG (codos 12, 13, 61) 12, 13, 61)	M.change KRAS (codon 12, 11 13, 61, 117,	Al change IRAF (codon 600)	AA.change PK2CA (codon 542, PTEN 545, 546, expressio	ADC index	NT	6.001pg/mi Cetux Cetux	a 0.1µg/mi Cetux	tygini Cetux	18µgimi Cotax	25µg/mi Cetux	Sõpgimi Cetux	100µg/mi Cetux
C18 C105 C125PM C145	ECACO ECACO ECACO	12022901 12022905 12022905 12022905	January 2013 January 2013 January 2013 February 2012	x y x x	11 10 11 11 14		12 12 9	10 8 9	12 10 10	12 4 54 11 12		11 12 12 9	20 21.2 22.2 28 29 22.2	4 9 4 7 9	10	9 11	14 15 18 17 19 14 18 15		MGI MGG MGG	WT WT WT WT WT WT WT WT	wT p.G12C p.K117N p.G13D	TW TW TW	WT positiv WT positiv WT positiv WT positiv	562.44 4972.23 1048.79 634.22	0	-2.57 -0.99 1.94 6.94 -4.71 8.22 0.14 4.94	-2.84 2.41 9.09 0.45	-1.41 12.45 9.19	-0.97 19.31 17.15 1.75	1.05 20.78 0.28 2.13	6.62 24.39 4.82 2.85	4.71 20.06 4.65 3.22
C179 C32 C79 C75	ECACO ECACO ECACO ECACO	10/ 12022908 12022909 12022907	Calculary 2013 January 2013 January 2013 January 2013	X Y X X X Y	12 10 11 12 12		10 11 50 8 12	10 9 9	12	4 11 4 11 12 10 13		9 12 12 9 11 12 9 11	29 22 2 28 20 29 21 2 21	7 9 a 5 6 9 a 9	2 4 2 50	9 9 11	18 99 54 58 17 58 17		MSI MSG MSG	WT WT WT p.G12D WT WT WT WT	GERD q TW TW TW TW	TW TW TW TW	WT postor WT postor WT postor WT postor WT	900 50 534 23 52933 35 52933 35	0000	-100 215 248 -676 602 505 11.58 42.74	2 M 4 N 19 10 55.59	2.40 -3.28 -44.51 -66.56	-1 09 -2 73 52 09 72.34	2 25 -4 12 -60 04 -77.90	5 29 7 41 67 78 77.62	6 90 1 50 60 81 74.34
C 60 C 64 C 69 CAC 02 CAR1	ECACC ECACC ECACC ATOC ATOC	12022402 12022402 12022402 12022402 12022402 12022402	January 2013 January 2013 Bafore 2015 Acci 2010		10 12 11 12 10 11 9 11	2	11 12 12 11 12 12 12 13 12	10 11 9 11	11 11 12	# 11 11 12 # 54 11 12 # 12	54	9 10 12 11 12 12 12	28 29 20.2 28 20 20 22 20 22			12 9 10 11	14 19 14 15 16 17 16 19 14 19		MEG MEG MEG MEG	WT WT WT WT WT WT WT WT WT WT	WT WT WT	TW WT TW TW	Vitaq TW Vitaq TW Vitaq TW Vitaq TW Vitaq TW	2006.00 2041.12 15079.23 2007.05 4854.64	0000	-0.66 0.45 8.79 10.71 11.68 34.54 4.25 12.51 -0.70 -2.74	1.00 21.44 40.70 10.00 5.00	2.01 23.47 58.43 13.54 19.52	6 20 27 20 64.00 15 20 15 50	4.44 37.18 72.40 11.60 22.83	15.03 36.78 76.18 15.94 24.29	24.44 23.64 73.64 54.05 29.19
CL11 CL14 CL34	JCR9 0.947 0.947 0.947 0.947	ACCAR ACCAR ACCER ACCER	March 2012 October 2010 October 2010 October 2010	× × ×	10 13 11 12 11 56	2	11 12 12 +	8 6 11 10	8 1 10.2 45 12 15	12 9 12 11 4 11 13	53	9 10 44 45 11 12 9 11	28 21 34 30 0 28 29 30	6 6 6 6 6 7	9 4 8	11	14 15 19 14 14 19 14 14 19		MGI MGG MGI	WT WT WT WT WT WT WT WT	Tw HRAD q WT WT	WT WT WT p V6006	vitace Tw vitace Tw vitace Tw vitace Tw	15413 17 4345 11 10895 69 2162.36	0 0 0	28 00 44 00 0 00 0 0 04 15 39 29 30 0.56 0 05	62.09 3.93 26.77 4.57	67.59 0.14 40.43 7.56	73 21 4 60 47.75 7.54	73 29 4 64 40 51 7.20	77 69 6 00 55 50 12 61	80 15 4 69 55 72 11.49
CL40 C0115 C0CMI C0GAI C0GAI	Diski Diskinger Diskinger Disk Linker Disk Linker	France Productor Austria Austria	Auty 2011 And 2010 November 2011 November 2011	X	10 12 11 12 12 11 10 11 9 11	2 13 14 1	11 90 12 11 12 11 12 11 15	9 7 7,1 12	9 16.2 11 10 12	# 11 12 55 10 # 11 14 15		11 12 10 11 12 12 9 13 11 12	20 22.2 34.2 25 29 21 27 34 30 20.2	0 5 9 9 7 5 9 9	a a a 9	9 10 11	16 17 15 20 16 17 16 18 14 19		MG MG MG MG MG	WT WT WT WT WT WT WT WT	TW TW TW TW TW	p VIOOE WT WT WT	Name Two Market Two Valace Two Valace Two Valace Two	4770.78 2196.48 16779.12 13201.17 11155.18	0000	0.59 6.07 2.75 1.98 18.12 28.69 -1.30 4.31 2.47 21.52	7.56 4.47 49.17 37.70 26.73	4.62 5.34 58.39 53.44 29.15	19.40 7.11 70.60 59.94 49.99	17.32 2.42 74.39 67.64 53.12	24.01 11.02 78.79 66.21 55.49	20.99 13.60 64.87 68.21 60.62
COGAE COGAE COGAE	Ord Liber Ord Liber Ord Liber	Austria Austria Austria Austria	November 2011 November 2011 November 2011 November 2011	X X X Y	10 11 10 12 11 10 11		12 16 12 13 13	11 8 10 10	13 11 111 11	11 12 16 12 13 12	2	11 12 13 32 11 11 12 11 12	29 362 302 312 29 29	4 9 7 5 4	a a a 11	10	18 20 21 17 19 18 20 18 20	1	MSG MSG MSG	WT WT WT WT WT WT WT WT	p.G129 p.G130 p.G130 p.G130	TW TW TW TW	Vitiag TW tun TW Vitiag TW Vitiag TW	7769 69 602 64 712 58 290 48	0 0 0 0	11 19 8 28 -0 51 -1 71 -1 03 -7 82 -2 69 -2 69	25 13 -1 80 -1 87 -1 87 -1 87 	24 15 -1 02 2 29 -2 05	28.99 -249 -442 -275	22.95 0.21 -1.58 1.24	27 75 4 79 9 12 2 54	45 03 2 75 0 90 1 59
COL0320 COL0320 COL0320	ATTOC ACTION ACTION ACTION	2003 2003 107 CCL-220	November 2011 Before 2015 November 2011 July 2011		11 12 11 11		9 9 9 9 9 9 9 9 9 9 9		11 12 12	11 12 11 12 11 12 11		12 13 12 13 11 12 11 12	202 202 202 202 202 202 202 202			9 9	15 12 15 12 15 12		MGG I	WT WT WT WT WT WT WT WT	Ter Tw Tw Tw Tw	p VSXG p VXXG WT WT	vision TH vision Th vision Th vision Th	100,04 2110,04 47,20 		1.14 2.45 -1.14 6.25 1.54 -1.20 -0.90 1.96	1.00 1.00 2.04 4.62	2.40 6.69 -1.27 -9.16	145 4 00 2 04 4.77	2.27 2.82 -1.41 -4.25	604 12.69 6.34 -3.22	124 0.49 156 4.58
COLOSSINGR COLOSSIN COLOSTS COLOSTS	DSM2	Gamani ACC194 USA	February 2012 December 2011 October 2010 Dafore 2010		11 15 12 11		12 9 16 13 14	9 7.2 8 12	12 10 12	11 10 12 2 11		11 12 9 12 9 12	31 202 29 202 27 29 29 222	6 9 9 9 7 5		9 10 11	15 18 15 56 17 16 19 17 18	7 18	MGG MGG MGG MGG MGG	WT WT WT WT WT WT In out door because of the m	wT p.A160T p.G12D elercens arisis of	WT WT WT Drig cell fine	WT postor WT postor WT postor	1006 72 -110 90 -718 07	0 0	-0.99 -0.24 1.46 0.41 0.23 -1.60	-1 66 4 02 0 57	-2.23 -0.40 -1.66	-2 50 -2 20 -6 69	-1 04 -1 90 -3 44	672 -069 -615	921 045 -349
CELOSER DIFI DLD1 FET	CLS CLS Dr Gausins ⁴ ATTO ISONP	Soain Concess Soain Concess Italy	December 2011 December 2010 December 2010 December 2012	X X X X X	11 94 11 12 10 11 11 12 12	2	11 12 11 12 11 12 13	10 10 10	12 12 11	11 4 11 8 11 11		11 12 12 12 13 12	27 28 29 20 22 29 32 2 28 32 2	7 9		9 9 11	15 174 17 19 17 18 18 19 16 17		MSG MSG MSI MSG	WT WT WT WT WT WT WT WT WT WT	wT wT p.G130 p.G130	WT WT WT	WT Date WT Date WT Date WT Date WT Date	915.22 1858.22 16123.06 98.34 1348.64	0 0 0	1.56 8.21 2.79 -148 1.29 41.41 0.54 1.48 -0.01 3.42	-0.03 -4.99 71.45 5.87 0.75	2.43 -1.64 78.26 0.96 2.21	1.69 1.63 80.82 	4.05 4.91 91.32 0.65 4.26	5.11 7.69 60.92 1.24 8.76	1.95 10.40 80.23 40.25 6.44
GEO GPID HCA24	ECACO ECACO ECACO ECACO ECACO	1146- 95290714 95200714 95200715 6051902 2101604	Extranse 2012 January 2012 January 2012 January 2012 October 2015	÷.	9 12 9 12 11		12 11 12 12 12 11 11 12	10 9 9	11 111 11 11 112 113	# 12 # 12 # 12 # 12 # 12 # 12		12 11 12 11 12 11 12 11 12 11 12	29 29 29 20 29 20 20 20 20 20 20 20 20 20 20 20 20 20	7 9 7 9 6 9		11 11	15 17 52 17 52 17 52 17 52 17		MG MG MG MG MG	WT WT WT WT WT WT WT WT WT WT	p.G12A p.G120 p.G120 wT wT	Tw Tw Tw Tw Tw	WT TW VERICE THORAT E VERICE THORAT E VERICE TW E VERICE TW E VERICE TW E VERICE TW E	460.44 5160.15 3031.60 5854.93	0000	478 -548 819 899 213 819 809 815 957 -114	4 79 17 22 19 64 12 79	-1 01 17 60 26 25 19 46	-4 05 24 91 29 99 24 27 60 05	-3 (9) 22 70 21 02 27 63 24 20	4 47 27 34 34 64 28 52 78 55	1 41 26 15 40 25 20 00 20 51
HCAP HCC3998 HCT115 HCT15	ECACC NORO NORO NORO	6001902 50N28 50N28 50N16	Before 2010 May 2011 May 2011 May 2011	X X X X X	11 54 11 13 7 90 12	2	9 11 12 12 12 11	11 10 11 10	12 11 12 12 12	9 11 12 10 12		10 11 9 12 11 12 12 13	26 28.2 29 20 29 20 29 20 29 2022	292 6 7 7 9 7 9		10 11	14 19 18 19 17 22 18 19		MSI MSI MSI	WT WT WT WT WT WT	wT p.A140T p.G13D p.G13D	TW TW TW TW	WT positiv WT positiv p.H1047R positiv p.5545K positiv	9721.99 925.45 1629.90 -434.22	0000	8.00 17.10 -2.41 -2.35 -0.46 7.54 1.92 2.54	19.77 -0.19 9.02 0.30	20.46 -1.20 9.09 -0.79	40.52 2.07 4.75 0.25	46.90 0.41 6.92 0.70	45.85 5.55 7.40 0.88	55.85 5.14 93.99 4.20
HOCI14 HDC114 HDC125 HDC142 HDC142	ATTO: 0wC23 0wC23 0wC23 0wC23	Germany Germany Germany Germany	November 2012 November 2010 November 2010 November 2010		12 11 9 11 11 12 11 12	2	12 12 10 14 12 12 12 12	10 10 10 8 9	11	11 12 9 4 9 12		12 13 11 13 12 12 12 13 12 13	29 20 2 28 20 26 28 27 28 20 22.2	2 9 9 9 9 9 7 5		11 8 11	10 10 14 19 16 19 17 19		MG MG MG MG MG	WT WT WT WT WT WT WT WT	TW TT TT TT TW	WT p V6006 WT WT	vitaoq TW vitaoq TW vitaoq TW vitaoq TW	1952 58 5223 58 1927 66 12203 91 -150.05	0000	-0.00 -2.68	-09 42 11 20 2 42 29 27 4.18	4 00 21 71 2 92 47.06 2.07	22 35 5 90 57.97 - 2.94	2 35 25 44 8 45 57.17 - 3.79	25 38 244 58 54 1.52	20 64 20 71 4 72 66.77 2 92
HDC54 HDC5 HDC52 HDC52	0xx23 0xx23 0xx23 0xx23 0xx23	Germany Germany Germany Germany	November 2010 November 2010 November 2010 November 2010	X Y X X X Y	11 12 13 10 12	2 13	12 12 11 11 12	8 8 9 7	9 10 12 10	11 12 11 9 11 11 12		11 12 12 12 13 14 15 15	30.2 32.2 25 29 30.2 26 31.2	9.3 9.3 7 6 9 3	9	10 10 9	17 19 15 19 17 19 18 19		165 165 165	WT WT WT WT WT WT WT WT	WT p.G12D WT WT	TW TW TW TW	višaog TW višaog TW višaog TW višaog TW	7716.15 1044.27 14156.92 1812.42	0000	6.14 54.31 -1.76 0.38 0.31 23.96 -0.25 5.49	26.04 0.51 46.65 6.00	21.80 0.40 56.07 6.25	27.72 1.06 43.27 # 48	25.53 0.75 71.55 0.91	25.71 5.61 71.34 12.34	45.55 8.62 71.69 5.91
HR425 HR0C18 HR0C24 HR0C32 HR0C32	Dr Linsbachar ^b Dr Linsbachar ^b Dr Linsbachar ^b Dr Linsbachar ^b	Germany Germany Germany Germany	January 2012 January 2012 September 2011 January 2012 January 2012	X Y X X	11 12 11 12 14 12 13	2	9 11 90 11 12	11 8 11 8 12	11 12 11	11 7 12 12 12 12		2 2 1 2	20 20.2 27 28 21 29	7 6 6 5 8 5 7 9		12 51	17 17 20 19 19 19 19 19 19 19 19 19 19 19 19 19		MEG MEG MEG	WT WT WT WT WT WT WT WT	p.A160V WT p.G12A p.A166T	WT p.V8006 WT WT	p ES45K positiv WT positiv WT positiv WT positiv	12362 48 3474 66 587.69 6995.90 2065.57	0000	2149 25.5 -0.05 -1.90 2.30 1.81 2.12 8.50 -6.21 4.48	2 10 2 49 57 37 16 26	6 20 1.51 20 26 17.42	#122 # 74 2.41 21.87 90.50	0 90 7.65 24.55 15.72	63 53 24 92 2 89 40 69 19 85	6211 1999 629 27.10 14.50
HROC46 HROC49 HROC40 HROC47	Dr Lissebacher ^A Dr Lissebacher ^A Dr Lissebacher ^A Dr Lissebacher ^A	Germany Germany Germany Germany	January 2012 January 2012 January 2012 Generatory 2013	X Y X Y X	12 10 11 12 12		13 12 13 12 13 14 11	8 10 9	11 12 11	12 11 12 4 10 14	9	12 13 12 12 10 11 12	28 22.2 29 20.2 20 22.2 28 29	6 9. 7 9. 6	a a a a 11	12 12 11 12	15 17 17 19 17 19 17 19		MEG MEG MEI	WT WT WT WT WT WT WT WT WT WT WT WT	p.G12V WT p.G12V WT	WT WT WT p V8006 WT	vidaoq TW vidaoq TW vidaoq TW vidaoq TW	1241.91 2749.48 -1710.42 2019.72	0000	6.78 -4.77 6.50 10.14 8.40 0.96 -1.59 1.63	4.01 16.02 -2.41 6.09	13.69 16.15 -7.38 7.15	4.71 54.70 4.47 8.78	12.91 17.76 -0.96 6.31	9.58 17.66 -7.83 10.26	-0.27 22.07 -12.29 12.99
HT28 HT55 HUTU80 KM12	ECACO ATOC NOMO	603336 85053105 1/72-40 502247	May 2011 January 2012 Dafree 2015 May 2011	X X X X	11 12 11 11 12 15 12		11 12 11 12 11 12 12 13	10 11 9	13	41 42 9 13 4 11 12 15		11 12 54 50 11	29 30 28 31 32.2 27	4 9 4 9 7 93	4 4 9	9 9 11 12	17 99 54 57 55 59 47 99		MGG MGG MGG	WT WT WT WT WT WT WT WT	TW TW TW	WT WT	Two votanog Two votanog Two votanog Two two Two two	1255.35 4342.82 -625.35 -771.60	0	-055 -259 -2.65 -0.24 1.11 -1.51 -1.55 -1.55	0 33 2.86 2.84 2.53	-0 76 12.38 -2.56 -4.30	-2 13 17.77 -4.20 -4.60	0 83 16.83 4.69 4.99	872 23.44 6.25 4.05	22.03
KM12C KM12L4 KM12SM KM20 LIM121S	KOLA KOLA KOLA	80015 80014 80017 80017 80017	November 2010 November 2010 November 2010 Sectomber 2010	***	10 12 12 15 12 11 12 16		10 11 11 11 11 12 11 12		1 1 1 1 1 1 1	12 19 10 19 10 19 11 4 10		11 11 52 11 52 11 52 11 52 6 6 41 43	27 34.2 27 34 27 34.2 29 30 34 31	93 93 4 93		12 12 12 9	17 16 17 16 17 17 19 17 19			WT WT WT WT WT WT WT WT	TW TW TW TW	WT WT P WXX6 WT	The Text of Te	-46.10 147.82 -109.42 -2210.47 10000.06	•	-4.99 -5.00 -7.19 -3.99 4.71 -4.45 -10.94 -7.29 04.64 -4.37	4 00 - 2 65 - 7 65 - 4 22 - 50 46	-6.25 -2.13 -9.60 -10.24 -59.05	4.12 	4.11 -2.49 -4.90 -14.25 -60.10	0.42 0.16 -0.25 -0.25 -0.25	248 490 7,77 4594 40,42
L1M2009 L1M2009 L1M2405 L1M2412 L1M2412	Dr. Walker Dr. Walker Dr. Walker Dr. Walker	Australia Australia Australia Australia Australia	December 2010 December 2010 December 2010 December 2010	X Y X Y X Y X	10 11 14 15 9	5	11 14 11 12 12 50	8 13 7 8	8 8 11	8 12 8 11 8 11 8 13		11 11 14 15 11 12	29 29 30 29 30 29 30 28 29 38 30	6 9 93 20 6	2 9 8 7	11 8 9	17 20 19 14 19 14 17 14 17	Ħ	453 453 453 453	WT WT WT WT WT WT WT WT WT WT WT WT	p.G12A p.G12C WT WT WT	WT p V8006 p V8006 p V8006	vidaoq TW vidaoq TW lun TW lun TW	1990 11 1968 03 962 68 -445.10 -803 68	0000	-4.91 -4.20 -0.15 -0.34 8.06 5.78 -2.97 5.11 -5.99 4.39	-2.69 1.78 8.34 1.22 -41.78	0.47 2.01 5.01 2.12	1 17 12 64 -0.02 -0.73 -4.45	7.26 4.45 5.59 -0.09	7 25 4 74 2 70 -1.19 4 00	12 89 6 98 1 25
L192550 L192551 L0V0 L51034	CBA CBA ATOC ATOC	C9A-019 C9A-0170 CCL-229 CQL-2459	September 2011 Generative 2011 July 2011 July 2011	X Y X Y X Y	12 10 11 11 12 11 12	5 2 14 2	12 12 12 16 11 12 12 12 19	9 9 93 83	10 11 10 11 12	12 15 # 10 # 11 12		9 14 11 9 12 9 11	29 212 29 202 29 212 29 212 29 20	8 5 6 9 9.3 7 7		51 9 11	14 19 17 19 17 18 17		MSI MSI MSI	WT WT WT WT WT WT WT WT	p.Q01K WT p.G13D p.A160T	WT p.V8006 WT WT	WT posto WT RedtH.q WT posto WT Vision	1506.88 113.22 2196.41 568.29	0000	9.91 19.70 9.50 2.51 -4.03 -2.00 0.75 -1.04	19.36 7.42 0.13 0.45	12.05 0.45 -2.45 1.08	54.77 -2.97 1.54 2.17	11.95 0.89 6.62 3.70	5.42 -1.27 14.27 0.98	5.95 5.28 11.55 5.22
LS122 LS1247 LS180 LS180 LS110 LS110	ATOC ATOC ATOC ATOC	CCL-255 CL-447 CL-447 CQL-447 CQL-447 CQL-2474 CQL-2474	July 2011 January 2011 January 2011 July 2011 July 2011	X X X Y	10 11 12 12 10 10 10 10	14	11 11 15 10 15	9 50 7.3 8	11 9.2 11	* 12 11 12 12 * 12	14	12 11 12 11 12 11 12 12 12	20 21 22 21 29 21 29 22 20 20	24 4		9 9	16 17 15 17 15 17 16 17 16 17 16 17		153 153 153 153	W1 W1 WT WT WT WT WT WT WT WT WT WT	p.0125 p.0120 p.0120 wT p.0120	art p.VI006 wT	WT pasty WT pasty WT pasty WT pasty	-1000 79 1075 79 4070 79 -670 45 1716 36	• • •	-174 -425 228 5.41 464 225 -022 -2.18 -1.12 1.77	-90 32 0 37 0 30 1.09 2.30	-4.66 7.94 -0.00 -4.00 4.57	2.28 1.77 -10.54 -2.22 -6.29	477 480 4.00 7.28	4 80 9 72 4 46 4 12 8 41	4.74 4.11 0.69 -0.59 30.37
MD-STB MP101 NC1H680 NC1H500	ATCC ATCC ATCC	90011801 USA CCL-254 CCL-254	January 2012 June 2012 May 2012 Salawary 2011 Navamber 2010	× × ×	11 12 11 12 11 12	2	90 90 90 90	4 10 8 9	45. 11.2 11 12	45. 44 8 11 9 11 8 12 11 22		11 12 13 9 11 12 11 10	29 91 29 32.2 29 31 29 32.2 30 33.2	a a 7 9 7 9 9 7		11	46 35 18 19 15 15 16 59 16 56 17	7 14	MGG MGG MGG MGG	WT WT WT WT WT WT WT WT WT WT	WT p.G130 p.G01K WT WT	p WIOOK WT WT WT	WD TW vidaoq X5452.q	4700 50 455 20 -660 50 56710 76 1024 06	0	-2.45 -4.79 -0.40 -0.90 0.56 1.70 -2.45 4.06 -0.99 0.641	-4.00 -1.96 -40.08 -2.66	-0.64 -0.54 -0.54 -0.54	4 61 4 20 4 20 4 62 4 62	2.04 0.42 82.92 4.69	7.40 2.38 4.65 63.42 4.97	44 44 2 27 -2.54 44 31 -2.64
NCINESS NCINTS NCINTS OUM523	ATOC ATOC ATOC ATOC	0014 CCL-251 CCL-252 JCR81022	Nincember 0010 July 2011 Bafore 3015 40209	X Y X X X	10 11 11 11 12 12	2	11 11 11 11	10 10 10	11 11 12	11 12 4 11 11 12 12	5 A	11 12 11 12 11 12 9 12 11	21 21 22 22 22 22 22 22 22 22 22 22 22 2	9 6 9 7 9	2 <u>8</u> 2 <u>8</u> 2 11	9 11 11	12 34 17 16 17 16 17 16 19 16		MGG MGG MGI MGG	WT WT WT WT WT WT WT WT	p.G120 WT p.G130 WT	WT WT WT p.VI006	vitaoq Tw vitaoq Tw vitaoq Tw luc Tw	2621 58 1450.05 -313.64 1577.07	0000	-1.06 1.12 2.94 8.72 -11.78 -4.76 6.76 2.54	1 90 	746 -7.77 -9.20 1.07	10 15 0.44 -13 02 -3.04	15 84 -7.12 -12 32 0.88	17.98 9.69 -1.20 12.91	20.40 15.72 8.11 6.49
OKCO2 OKCO2 OKCO3 RC91 RC91	Deef Casuadoint Deef Casuadoint Deef Casuadoint HSRSS Deef Bart Venalataint	107 UK 107 20586256 1154	November 2010 November 2010 November 2010 March 2010 Austral 2010	X Y X X	12 9 12 12 10 12 4 11	2	12 11 12 11 11	8 10 10	12 12 11 12 12 14 15	11 10 11 11 12 10 12 4 11		11 12 11 12 9 11 12 11 12	20 20 28 29 29 30 30			10 11	16 17 17 22 16 19 20		MSG MSG MSG MSG	WT WT WT WT WT WT WT WT WT WT	W1 WT p.Q61H p.G12V WT	WT WT P VIOCE	WT posto- WT posto- WT posto- WT posto- p.H1047R posto-	654.50 57470.11 1076.96 2060.92 99.57	0000	-7.42 -4.53 24.42 -41.54 1.22 -4.51 1.45 -4.42 6.51 -3.66	-7.11 66.27 2.97 2.99 0.72	-6.98 78.30 4.54 6.01	4.16 82.85 8.96 90.99 4.16	-0.71 -00.16 5.20 12.00 3.31	-0.44 86.17 2.01 16.51 0.36	95.05 89.99 9.75 21.42 -2.64
RW7212 SKC01 SNU1022 SNU1042	Nem Arann ^a ATV: KCLB KCLB	Soain una.16 718 1560	October 2012 kdv 5011 November 2010 November 2010	X Y X X X Y	11 13 19 12		11 4 4 12 10 12	10 6 11 11	45. 12 13	10 12 11 12		a 10 12 9 9	20 21 34 55 29 20 28 29	6 9 22 6 5	11 4 11 8		18 44 44 17 19 17 19 20	0	MGG MGG MGI	WT WT WT WT WT WT WT WT	p.G120 p.G12V p.G12D WT	TW TW TW TW	vitace TW vitace TW vitace TW vitace TW	2990 63 	0 0 0	2 00 7 59 .40 /06 .45 30 -2.04 -2.09 17.50 12.43	7.45 0.49 0.09 21.45	11 64 -7 41 2.06 22.59	12.50 4.51 27.56	54 72 3 74 3 02 21.95	15.81 .7.90 1.21 23.65	14 34 0 64 5 55 32.57
SNU1047 SNU1181 SNU1235 SNU1406 SNU1411	KCLR KCLR KCLR	1992 1981 1995 1995 1995	Kincambar 0040 July 2012 Luke 0040 July 2012 Luke 0040	X X X X	9 11 12 12 11 12 10 12	2	8 12 90 12 11 12 11 13	8 10 8 8 11	10 11 12 11	10 11 10 11 12 4 9 10		9 10 12 12 11 12 12 13	29 202 29 202 30 202 30 202 30 202	9 9 9 6 9		9 11	14 21 14 12 17 15 12 20		MSI MSG MSG MSG	WT WT WT WT WT WT WT WT	p.G129 WT WT p.G12C	WT p.V6006 WT WT	Two	-1327.24 2977.62 -2184.11 * 783.28	0000	-6.29 -3.12 7.44 27.15 -0.14 -2.45 -2.76 -2.04	4.50 5.25 0.45	-7.13 8.20 0.50	4.55 11.50 9.24 -2.19	4.17 7.23 4.64	-6.49 18.44 -11.91 4.92	4.25 14.29 -18.04
SNU1460 SNU1544 SNU1684 SNU1746 SNU1746	KCLB KCLB KCLB	1600 1544 1984 1746 175	July 2012 kidu 0040 July 2012 kidu 0040 kidu 0040	X X X Y	12 11 12 9 50 7 9		11 12 90 11 11 12 12	8 11 9 14 8	11 12 13 11	9 12 9 11 8 9		10 8 14 9 10 9 14 9 10	20 2 21 29 20 28 20 28 20 30 22 3	4 7 7 5 9 6		11 10 9	14 18 54 20 21 54 17 58 54		MG MG MG MG	WT WT WT WT WT WT WT WT WT WT	WT WT p.G130 p.G130 WT	TW TW TW TW TW	witacq Tw vidacq RfabitH.q witacq Tw witacq Tw witacq Tw	12054 06 2506 72 7154 92 -2505 53 5176 **	0000	2 81 3 72 2 52 2 34 16 51 13 30 11 45 17 59 2 05 14 44	16.60 0.92 17.71 1.72	48 23 0.04 21.18 0.42 99.76	53 98 2.48 32.95 -16.35 22.47	58 27 7.12 29.42 -18.02 34.77	57 75 13 29 27 77 4 94 95 53	47 51 19.10 28.54 -12.65 39.59
5N2254 5N2253 5N2457 5N2477 5N2479	HOLD HOLD HOLD	115 254 638 479	November 2010 Lido 5015 November 2010 October 2012 July 2012	X X X X	10 11 10 11 11 12 10 11 10 11		9 9 10	11 13 11 8 9	10	4 12 11 8 11 8 10 9 11		9 12 9 12 11 12 9 11	24 20 20 20 20 20 20 20 20 20 20 20 20 20	7 5	8 8 8 7 9	12 11 8 12	14 17 14 17 14 19 14 19 14 19		MEG MEG MEG	WT WT WT WT WT WT WT WT	p.G120 WT p.G120 WT	TW TW TW TW	p.545A positiv WT positiv p.H1047R positiv WT null	11/0.45 1417.81 10221 76 4252 95 5044 47	0000	4.03 -2.93 14.99 -22.03 -0.28 -147 -24.90 -14.93	54.21 52.26 7.41 45.50	-1.91 -1.91 -14 43 -14 23 	2 18 41 51 15 01 42 04	49.17 49.20 20.15 37.70	825 51.53 20.97 53.56	22.50 54.37 24.13 40.90
5NU582 5NU581 5NU7890 5NU81 5NU81	KCLB KCLB KCLB KCLB	500 44 660 85 677	October 2012 November 2010 November 2010 October 2012 July 0110	X X X X X	11 12 11 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 14		90 111 54 90 111 11 40	10 8 11.3 52 60	11	10 8 9 12 9 10 11		11 12 12 12 9 12 10	20 29 29 21 20 21 20 21 29	9 9 7 6 7 5 7		11 9 11	14 17 16 18 16 17 16 17 16 17		MEG MEG MEG MEG	WT WT WT WT WT WT WT WT WT WT WT WT	WT p.G12D WT p.A160T WT	TW WT TW TW TW	vGaog TW vGaog TW vGaog TW vGaog TW vGaog TW	10216-08 2005.72 11134.12 659.04 0668.04	0000	12 32 22 72 -2 40 -0.44 -4.44 -4.00 -0.92 -2.35 -3.19 -7.75	29 75 0.96 4.44 -0.43 -0.73	27 89 5.04 4.06 -2.06 2.00	43.13 3.17 23.93 -1.10 4.30	46.83 5.59 55.25 -1.19 6.72	54 14 11.27 58.14 6.26 18.9 ⁶	51 75 12 80 58 52 2.11 15 45
SNUC1 SNUC2A SNUC2B SNUC4	XCLR XCLR XCLR	0000C3A 0000C3A 0000C3B 0000C4	November 2010 November 2010 November 2010 November 2010	X X X X Y	12 11 54 11 54	4	90 8 9 8 9 8 9	10 13 12 73	11 16 16 11	9 11 9 12 9 50 12 8 9		9 11 8 9 9 9 9	20 20 20	7 5 9 9 7 5	11 2 2 11	12	14 19 15 19 15 19 14 15		M69 M61 M61 M61 M61	WT WT WT WT WT WT WT WT	WT p.G120 p.G120 WT	WT WT WT WT	WT posty WT posty WT posty posty posty nut	1451.74 - 560.84 2001 17 9296 57	0000	220 574 0.20 4.97 -0.41 1.29 -9.74 7.21	7.47 1.98 5.80 19.89	5.36 1.16 4.65 22.14	\$75 3.41 11.92 35.35	2.37 1.31 13.14 44.69	8.68 -2.73 16.61 -48.32	11.71 7.28 19.25 47.80
SNUCS SW1115 SW1222 SW1417 SW1462	ATOC Or Method ATOC ATOC	00005 001-223 1154 001-228	May 2011 May 2011 May 2011 May 2011 May 2011		11 12 10 11 11 12 11 +2		11 12 11 12 9 12 12 13 14	12 12 10 9	12 11	11 54 9 12 12 13		8 13 9 12 9 10 11 12 11	29 312 28 29 29 322 29 322 30 312	6 5 6 202 6 6 6 7	8 8 90 8 8	11	15 14 19 18 15 19		MGG MGG MGG MGG	W1 W1 W7 W7	#1 p.G12A p.A140V WT p.G12C	wr wr p vector p vector wr	Vision Two vision Two line Two vision Two vision Two	542 73 228.80 1308.71 1133.22 2042.14	0000	-0.40 -1.84 -5.54 -5.11 -0.10 3.22 -0.65 0.19 -0.421 A 07	-5.19 -7.71 1.35 -0.41 	-0.83 -3.47 5.25 -0.58 1.68	472 4.57 5.34 -1.27 11,22	-0.69 -1.99 -6.47 -1.94 -0.29	4 94 2.00 7.09 8.32 13.19	4 72 2 42 4 11 4 71 12 19
5W402 5W43 5W400 5W600	ATOC ATOC ATOC ATOC	CCL-220 CCL-221 CCL-221 S07424	40034 March 2011 January 2011 May 2011	× × ×	10 12 9 10 12 54 13 54		11 90 14 12 12	9 9 8	* 10 *	10 11 12 12 12		10 12 11 12 12 9 12	28 29 28 31 30 302 30 302	4 4 4	4 6 11	93	14 18 18 20 16 16		MSG MSG MSG	WT WT WT WT WT WT WT WT	p.G12V WT p.G12V p.G12V p.G12C	Tw Tw Tw Tw	p QS46K positiv WT positiv WT positiv WT positiv WT positiv	4077.36 1219.28 1825.71 -362.57	0000	-228 -141 -666 -477 -682 -160 -084 -328	4.22 6.27 2.54 -4.86	7 69 0 50 2 90 -4 96	11.45 6 21 1 20 4 77	17.40 7.12 5.60 4.12	18.75 5.22 2.42 0.25	27.30 4.66 8.18 2.68
200227 SW948 T84 V411 V457	ATCC ATCC Dod Machaultr Dod Machaultr	CCL-227 CCL-227 USA USA	40034 40024 kdc 2011 March 2012 March 2012	x x x	12 19 11 12 11 12	2	11 12 12 11 11 16	9 9 9 9 9 9 12	11 10 11 11	12 10 11 9 9 9 12 11 12		11 12 10 11 11 12 11 12 11 12	25.2 29 25.1 29 25 29 22.2 29 22.2 29 22	4 9 6 5 9 7 5		11 11 9	16 19 16 18 17 18 15 18		MEG MEG MEG	WT WT WT WT WT WT WT WT	p.Q41L p.Q13D p.Q13D p.Q13D p.Q12D	TW TW TW TW	ville, vilace X542.c vilace X542.c Vilace TW Nace TW	1859 80 51.08 2226 70 8 807 71 -099 99	0 0 0 0	-0.91 0.66 -0.91 0.66 -1.11 -0.07 1.16 0.18 5.71 2.31	0.47 0.47 2.24 2.50 4.41	= 45 1.08 5.29 -1.90 2.24	0.06 0.06 5.14 2.74 0.29	0.92 0.96 2.84 4.57	0.07 15.59 4.46 -0.28	1.38 1.38 1.38 1.38 1.30 1.30 1.30 1.30 1.30 1.30 1.30 1.30
V481 V702 VAC0422 VAC05 VAC05	Deel Markovitr' Deel Markovitr' Deel Markovitr' Deel Markovitr' Deel Markovitr'	USA USA USA USA	March 2012 March 2010 Autout 2010 Celevany 2012 February 2012	X X X X X X	7 12 9 12 9 10 11 12 11	2	8 11 10 9 10 15 11	8 11 59 11 50 7	12	11 12 9 12 11 12 9 12 9 12		11 15 9 11 14 9 12 9 12	21 29-2 21 28 23-2 24-2 29 29 21-2	7 9 9 9 9 9 9 9 9		9 10 8 11 10	17 19 21 19 21 14 19 14 19 14 19 19 20 22	2	19 19 19 19 19	w7 W7	p.G12C p.G12D WT WT	WT WT p.V600E p.V600E p.V600E	p (2646P null WT positiv- WT positiv- p.H1547R positiv- WT positiv-	-2706.95 2291.68 1876.61 1544.19 3625.00	0000	-3.90 -4.88 0.25 4.27 0.50 4.25 7.12 3.24 22.28 -1.05	-7.47 6.91 10.97 1.72 11.78	-8.04 6.51 9.55 4.74 16.15	-10.72 7.65 15.00 6.94 20.34	-10.22 9.04 14.11 0.55 14.28	-13.27 11.90 9.25 7.99 18.95	-17.06 12.41 5.09 8.42 19.21

Supplementary CRC cell lines fro Benjamini-Hockb	Data 2. Classific om GSE59857 we erg (BH)-correcte	cation of 151 CR are assigned by the d FDR. In the BH	C cell lines into ne NTP algorithm I.FDR columns, g	o molecular sul to a molecular grey shading inc	btypes accordin subtype for each licates low class	ng to five transcr of five different cl ification confidenc	riptonal classifi assifiers. For ea e (FDR > 0.2).	iers. ach classifier, in ac	ddition to the sub	type class, the ta	ble reports the
GEO ID	Cell Line	Sadanandam class	BH.FDR	Marisa class	BH.FDR	Budinska Class	BH.FDR	De Sousa E Melo class	BH.FDR	Roepman Class	BH.FDR
GSM1448103	COGA12 HDC54	Enterocyte	0.000990	C6	0.001235	Subtype B	0.001794	CCS1	0.030964	B-Type B-Type	0.020127
GSM1448175	DIFI	Enterocyte	0.000990	C6	0.000751	Subtype A	0.001794	CCS1	0.002580	B-Type	0.003871
GSM1448185 GSM1448199	HDC143 C32	Enterocyte Enterocyte	0.000990	C6 C6	0.000751	Subtype B Subtype E	0.185241 0.005636	CCS3 CCS1	0.401371 0.002580	B-Type B-Type	0.252024 0.003871
GSM1448208 GSM1448215	SNU479 NCIH684	Enterocyte	0.000990	C6 C5	0.000751	Subtype A Subtype C	0.001776	CCS1	0.002580	B-Type B-Type	0.006136
GSM1448110	V481	Enterocyte	0.002293	C3	0.000751	Subtype C	0.001510	CCS2	0.392312	A-Type	0.169639
GSM1448221 GSM1448205	V411 SNU1684	Enterocyte Enterocyte	0.004915 0.007631	C6 C6	0.002475	Subtype B Subtype D	0.201545	CCS1 CCS2	0.006773	B-Type C-Type	0.779002 0.061305
GSM1448157	SNU1411	Enterocyte	0.013285	C3	0.010619	Subtype E	0.762087	CCS3	0.910059	C-Type	0.992003
GSM1448189	SNU1040	Enterocyte	0.024998	C6	0.002706	Subtype D	0.004473	CCS3	0.279400	C-Type	0.127088
GSM1448220 GSM1448151	VACO6 COGA5L	Enterocyte Enterocyte	0.025762	C3 C5	0.016990	Subtype C Subtype B	0.001510	CCS2 CCS1	0.008386	A-Type B-Type	0.003871
GSM1448080	HROC69	Enterocyte	0.048097	C6	0.017273	Subtype B	0.217986	CCS1	0.191832	B-Type	0.159829
GSM1448127 GSM1448099	COGA1	Enterocyte	0.059593	C2 C2	0.001235	Subtype B Subtype E	0.036904	CCS2	0.004328	A-Type A-Type	0.232881
GSM1448111 GSM1448203	HT115 C84	Enterocyte	0.149772	C2	0.340635	Subtype C Subtype E	0.001776	CCS2 CCS1	0.088469	A-Type B-Type	0.059678
GSM1448202	C80	Goblet	0.000936	C3	0.014100	Subtype B	0.002157	CCS1	0.040515	B-Type	0.110592
GSM1448070 GSM1448071	HDC9 HROC24	Goblet Goblet	0.000990	C3 C3	0.000751	Subtype A Subtype C	0.002922	CCS2 CCS2	0.057591	C-Type A-Type	0.174684
GSM1448087	SNUC2B	Goblet	0.000990	C2	0.000751	Subtype C	0.006194	CCS2	0.048453	A-Type	0.054044
GSM1448164 GSM1448193	LS1/41 SKCO1	Goblet	0.000990	C3 C3	0.000751	Subtype A Subtype A	0.001776	CCS2 CCS1	0.859147	A-Type C-Type	0.044358
GSM1448206	SNU1746	Goblet	0.000990	C3	0.000751	Subtype C	0.001776	CCS2	0.005920	A-Type	0.784579
GSM1448141	LS513	Goblet	0.001267	C3	0.000751	Subtype C	0.090861	CCS1	0.300701	B-Type	0.873060
GSM1448119 GSM1448139	CL34 LS180	Goblet Goblet	0.006785	C2 C2	0.000751	Subtype C Subtype A	0.008933	CCS2 CCS2	0.009710	A-Type C-Type	0.293298
GSM1448188	KM12SM	Goblet	0.009535	C2	0.011895	Subtype A	0.231325	CCS2	0.011940	A-Type	0.272147
GSM1448120 GSM1448073	KM12	Goblet	0.010379	C3 C3	0.006564	Subtype C Subtype A	0.072921	CCS2 CCS2	0.450579	A-Type A-Type	0.484628
GSM1448186	KM12C	Goblet	0.012324	C3	0.005730	Subtype A	0.086067	CCS2	0.303117	C-Type	0.337246
GSM1448094 GSM1448159	SNU81	Goblet	0.019742	C3	0.025757	Subtype C	0.001776	CCS3	0.764465	A-Type A-Type	0.036089
GSM1448187 GSM1448204	KM12L4 C99	Goblet Goblet	0.090582	C2 C3	0.005921	Subtype A Subtype A	0.637688	CCS2 CCS1	0.600230	A-Type B-Type	0.747456
GSM1448217	SNU407	Goblet	0.161949	C2	0.000751	Subtype C	0.028037	CCS2	0.162952	A-Type	0.635107
GSM1448136 GSM1448218	SNU769B	Goblet	0.162815	C1 C3	0.517844 0.029487	Subtype A Subtype C	0.273869 0.001794	CCS2 CCS1	0.446369 0.105451	B-Type A-Type	0.826927 0.006136
GSM1448081	V457	Goblet	0.332352	C2	0.001006	Subtype A	0.487231	CCS2	0.026638	A-Type B. Type	0.033065
GSM1448171 GSM1448172	COLO205	Goblet	0.350034	C3	0.386958	Subtype E	0.491730	CCS2	0.876460	С-Туре	0.339217
GSM1448179 GSM1448198	HT29 C146	Goblet Goblet	0.474598	C1 C2	0.516591	Subtype C Subtype E	0.001510	CCS3 CCS3	0.073678	A-Type C-Type	0.424097
GSM1448169	HCC2998	Goblet	0.709348	C2	0.720713	Subtype B	0.771782	CCS1	0.970676	A-Type	0.881046
GSM1448124 GSM1448135	VACO432 LIM2537	Inflammatory	0.817685	C2 C2	0.000751 0.000729	Subtype C Subtype C	0.051803	CCS2 CCS2	0.005134 0.004328	A-Type A-Type	0.072679 0.014506
GSM1448149	CL40	Inflammatory	0.000936	C6	0.000751	Subtype A	0.455554	CCS1	0.106044	C-Type	0.966413
GSM1448072	HROC87	Inflammatory	0.000990	C2	0.000751	Subtype C	0.002922	CCS2	0.009710	A-Type A-Type	0.059054
GSM1448086 GSM1448090	SNU61 T84	Inflammatory Inflammatory	0.000990	C6 C2	0.012015	Subtype C Subtype C	0.001776	CCS2 CCS1	0.214830	A-Type B-Type	0.126305
GSM1448095	HROC18	Inflammatory	0.000990	C1	0.002475	Subtype C	0.001776	CCS1	0.002580	B-Type	0.054090
GSM1448108 GSM1448118	LIM1215	Inflammatory	0.000990	C2 C2	0.000751	Subtype E Subtype C	0.448469	CCS1 CCS2	0.213236	A-Type	0.445882 0.103811
GSM1448138	LS123	Inflammatory Inflammatory	0.000990	C4	0.000751	Subtype C	0.001776	CCS3	0.134671	C-Type	0.015615
GSM1448158	SNU1460	Inflammatory	0.000990	C6	0.000729	Subtype C	0.002581	CCS1	0.002580	B-Type	0.059054
GSM1448195 GSM1448091	SW837 WIDR	Inflammatory Inflammatory	0.002593	C5 C2	0.019594	Subtype C Subtype E	0.001776	CCS3 CCS3	0.572890	C-Type A-Type	0.401727
GSM1448150	COGA5	Inflammatory	0.003907	C5	0.003195	Subtype C	0.001776	CCS3	0.268213	C-Type	0.383461
GSM1448131 GSM1448190	SNU1235	Inflammatory	0.006519	C2 C4	0.000751	Subtype D Subtype D	0.150440	CCS2 CCS3	0.218617	C-Type C-Type	0.082238
GSM1448168	LIM2550	Inflammatory Inflammatory	0.009365	C2	0.005536	Subtype C	0.001776	CCS2	0.002580	A-Type	0.003659
GSM1448145	NCIH747	Inflammatory	0.022430	C4	0.174720	Subtype D	0.209861	CCS3	0.789212	C-Type	0.078209
GSM1448163 GSM1448076	LOVO LIM2551	Inflammatory Inflammatory	0.024497	C2 C2	0.001735 0.000751	Subtype E Subtype C	0.051419 0.001776	CCS2 CCS2	0.089148 0.015055	A-Type A-Type	0.818343
GSM1448074	KM20	Inflammatory	0.044482	C5	0.005162	Subtype C	0.001776	CCS2	0.031017	A-Type	0.062896
GSM1448107	C170	Inflammatory	0.000717	C2	0.152224	Subtype D	0.264851	CCS2	0.073403	A-Type A-Type	0.421550
GSM1448067 GSM1448075	CO115 LIM2412	Inflammatory Inflammatory	0.137821	C2 C2	0.000751	Subtype C Subtype F	0.008933	CCS2 CCS3	0.002580	A-Type C-Type	0.003871
GSM1448174	COLO678	Inflammatory	0.319792	C4	0.054298	Subtype D	0.099640	CCS3	0.471947	B-Type	0.097957
GSM1448132 GSM1448165	SW48	Inflammatory	0.723828	C2 C2	0.002249	Subtype E Subtype C	0.061840	CCS2	0.002580	A-Type	0.361216
GSM1448178	HCT15 MIP101	Inflammatory Inflammatory	0.880083	C2	0.012357	Subtype C Subtype B	0.324066	CCS2	0.002580	A-Type	0.334894
GSM1448140	LS411N	Inflammatory	0.912018	C2	0.001235	Subtype B	0.697417	CCS2	0.002425	A-Type	0.449418
GSM1448089 GSM1448197	SW1417 C106	TA TA	0.000936	C5 C1	0.000751	Subtype C Subtype E	0.001510	CCS1 CCS1	0.116897	B-Type B-Type	0.004645
GSM1448079	HROC32	TA TA	0.000990	C1	0.000751	Subtype B	0.001510	CCS1	0.002580	B-Type	0.003871
GSM1448082 GSM1448083	HROC39	TA	0.000990	C5	0.000751	Subtype B Subtype B	0.001510	CCS1 CCS1	0.002580	B-Type B-Type	0.003871
GSM1448088 GSM1448101	SW1222 COGA3	TA TA	0.000990	C1 C2	0.001006	Subtype B Subtype B	0.001776	CCS1	0.002580	B-Type B-Type	0.003871
GSM1448109	HCA24	ТА	0.000990	C3	0.009407	Subtype B	0.018051	CCS1	0.002425	A-Type	0.205840
GSM1448112 GSM1448113	LS1034 SW1463	TA TA	0.000990	C1 C5	0.000751	Subtype B Subtype B	0.173422	CCS1 CCS1	0.002580	B-Type B-Type	0.006136 0.003871
GSM1448115	SNUC1	TA	0.000990	C5	0.000751	Subtype B	0.004473	CCS1	0.002580	B-Type	0.003871
GSM1448117 GSM1448128	HRA16 HDC82	TA	0.000990	C5 C5	0.000751	Subtype B Subtype B	0.006710	CCS1 CCS1	0.002580	B-Type	0.003871
GSM1448129 GSM1448142	HDC142		0.000990	C5	0.000751	Subtype B	0.002157	CCS1	0.002580	B-Type B-Type	0.003871
GSM1448146	OXCO2	TA	0.000990	C6	0.118519	Subtype B	0.001510	CCS1	0.002580	B-Type	0.003871
GSM1448153 GSM1448156	COLO94H RW7213	TA TA	0.000990	C5 C3	0.000751	Subtype B Subtype A	0.001776	CCS1 CCS1	0.002425	B-Type B-Type	0.004645
GSM1448167	COCM1	TA	0.000990	C6	0.000751	Subtype B	0.001776	CCS1	0.002580	B-Type	0.003871
GSM1448171 GSM1448176	GEO	TA	0.000990	C5 C3	0.000729	Subtype A Subtype B	0.380716	CCS1 CCS1	0.002580	в-туре B-Type	0.005514
GSM1448177	HCA46	TA	0.000990	C1	0.000751	Subtype A	0.009479	CCS1	0.002580	B-Type	0.003871

GSM1448191	SNU1406	TA	0.000990	C5	0.000751	Subtype B	0.001776	CCS1	0.009710	B-Type	0.003871
GSM1448194	SW403	TA	0.000990	C5	0.000751	Subtype B	0.001776	CCS1	0.002580	B-Type	0.003871
GSM1448200	C70	TA	0.000990	C1	0.051278	Subtype B	0.001794	CCS1	0.002580	B-Type	0.017056
GSM1448201	C75	TA	0.000990	C1	0.231460	Subtype B	0.001794	CCS1	0.002580	B-Type	0.007575
GSM1448213	SW948	TA	0.000990	C5	0.000751	Subtype B	0.006416	CCS1	0.116036	B-Type	0.013291
GSM1448106	Gp5D	TA	0.001267	C2	0.010018	Subtype B	0.002922	CCS1	0.002580	A-Type	0.159254
GSM1448114	SW1116	TA	0.001267	C5	0.000751	Subtype B	0.149920	CCS1	0.002580	B-Type	0.005514
GSM1448219	SNU1181	TA	0.001267	C5	0.000751	Subtype E	0.872426	CCS1	0.447968	A-Type	0.330043
GSM1448123	SNU283	TA	0.001610	C1	0.019548	Subtype C	0.001510	CCS1	0.002580	B-Type	0.010375
GSM1448100	COGA2	TA	0.002293	C5	0.023526	Subtype C	0.022248	CCS2	0.143282	B-Type	0.023586
GSM1448097	CCK81	TA	0.003595	C1	0.018872	Subtype B	0.001776	CCS2	0.002425	B-Type	0.677054
GSM1448214	NCIH498	TA	0.008861	C5	0.000751	Subtype B	0.748810	CCS1	0.513663	B-Type	0.003871
GSM1448098	RCM1	TA	0.075741	C1	0.000751	Subtype E	0.400633	CCS1	0.270296	B-Type	0.025545
GSM1448085	SNU1047	TA	0.096473	C2	0.000751	Subtype D	0.020879	CCS2	0.030964	C-Type	0.034726
GSM1448068	COLO201	TA	0.149543	C5	0.000751	Subtype C	0.001776	CCS2	0.053385	C-Type	0.339531
GSM1448105	GP2D	TA	0.153310	C2	0.001471	Subtype B	0.013420	CCS1	0.013418	A-Type	0.158788
GSM1448142	NCIH508	TA	0.219985	C1	0.121106	Subtype A	0.005414	CCS1	0.002580	B-Type	0.791861
GSM1448212	SW620	TA	0.280308	C5	0.000751	Subtype D	0.005414	CCS3	0.460303	C-Type	0.022363
GSM1448084	HROC46	Stem	0.000936	C5	0.000751	Subtype D	0.060681	CCS1	0.160733	B-Type	0.010375
GSM1448160	CACO2	Stem	0.000936	C4	0.002002	Subtype D	0.002922	CCS1	0.270967	B-Type	0.003659
GSM1448166	SW480	Stem	0.000936	C4	0.000751	Subtype D	0.001510	CCS3	0.021943	C-Type	0.007575
GSM1448077	OUMS23	Stem	0.000990	C4	0.000751	Subtype D	0.001510	CCS3	0.008386	C-Type	0.003871
GSM1448078	SNU1033	Stem	0.000990	C4	0.008071	Subtype D	0.007686	CCS3	0.015055	C-Type	0.006709
GSM1448096	SNUC2A	Stem	0.000990	C4	0.000751	Subtype D	0.005139	CCS3	0.014727	C-Type	0.020454
GSM1448122	MDST8	Stem	0.000990	C4	0.000751	Subtype D	0.001776	CCS3	0.002580	C-Type	0.003871
GSM1448134	OXCO1	Stem	0.000990	C4	0.000751	Subtype D	0.001776	CCS3	0.002580	C-Type	0.006136
GSM1448137	LIM2099	Stem	0.000990	C4	0.000751	Subtype D	0.001776	CCS3	0.002425	C-Type	0.003871
GSM1448170	CAR1	Stem	0.000990	C4	0.000751	Subtype E	0.001794	CCS3	0.029471	C-Type	0.004645
GSM1448180	HUTU80	Stem	0.000990	C4	0.000751	Subtype E	0.011042	CCS3	0.086656	C-Type	0.006136
GSM1448196	C10	Stem	0.000990	C4	0.000751	Subtype D	0.001776	CCS3	0.002425	C-Type	0.020454
GSM1448207	SNU254	Stem	0.000990	C4	0.000751	Subtype D	0.005636	CCS3	0.047736	C-Type	0.010960
GSM1448209	SNU503	Stem	0.000990	C4	0.000751	Subtype D	0.002581	CCS3	0.079715	C-Type	0.015007
GSM1448211	SNU977	Stem	0.000990	C4	0.001006	Subtype D	0.003252	CCS3	0.116734	C-Type	0.014506
GSM1448102	COGA8	Stem	0.001267	C1	0.001235	Subtype D	0.011042	CCS3	0.048430	C-Type	0.029072
GSM1448173	COLO320DM	Stem	0.001267	C4	0.000751	Subtype D	0.001510	CCS3	0.005134	C-Type	0.291489
GSM1448144	NCIH716	Stem	0.001973	C4	0.005337	Subtype D	0.002157	CCS3	0.113874	C-Type	0.158788
GSM1448152	COLO320	Stem	0.002293	C4	0.000751	Subtype D	0.001776	CCS3	0.014727	C-Type	0.051607
GSM1448125	HCA7	Stem	0.002593	C2	0.000751	Subtype D	0.022645	CCS2	0.032419	C-Type	0.011612
GSM1448148	CL11	Stem	0.006247	C4	0.000729	Subtype D	0.007686	CCS3	0.032419	C-Type	0.038077
GSM1448104	HDC8	Stem	0.015689	C4	0.001006	Subtype C	0.132201	CCS3	0.270032	A-Type	0.159419
GSM1448093	HROC80	Stem	0.042435	C2	0.024357	Subtype C	0.023968	CCS3	0.462838	A-Type	0.204528
GSM1448162	HCT8	Stem	0.069052	C4	0.003407	Subtype D	0.011535	CCS2	0.189027	C-Type	0.003659
GSM1448182	COLO320HSR	Stem	0.072001	C4	0.000751	Subtype D	0.004473	CCS3	0.017030	C-Type	0.074093
GSM1448192	SNU1544	Stem	0.080769	C2	0.001235	Subtype D	0.301940	CCS2	0.301107	A-Type	0.164496
GSM1448092	HCT116	Stem	0.162033	C4	0.005162	Subtype D	0.067002	CCS3	0.115157	C-Type	0.004645
GSM1448161	DLD1	Stem	0.311486	C4	0.072692	Subtype D	0.445976	CCS3	0.138248	C-Type	0.450929
GSM1448126	LIM2405	Stem	0.370679	C2	0.001235	Subtype D	0.051330	CCS2	0.189831	C-Type	0.068408

Table lists primers used to sequence 16 hot spot mutations in KRAS, NRAS, HRAS, BRAF and PIK3CA.										
Gene/Exon	Primers (5'-> 3')									
KRAS EX2	FORW	GGTGGAGTATTTGATAGTGTATTAACC								
KRAS EX2	REV	AGAATGGTCCTGCACCAGTAA								
KRAS EX2	SEQ FOR	TCATTATTTTTATTATAAGGCCTGCTG								
KRAS EX3	FORW	AAAGGTGCACTGTAATAATCCAGAC								
KRAS EX3	REV	ATGCATGGCATTAGCAAAGA								
KRAS EX3	SEQ FOR	CCAGACTGTGTTTCTCCCTTC								
KRAS EX4	FORW + SEQ	TGGACAGGTTTTGAAAGATATTTG								
KRAS EX4	REV	ATTAAGAAGCAATGCCCTCTCAAG								
NRAS exon2	FORW + SEQ	GTACTGTAGATGTGGCTCGC								
NRAS exon2	REV	AGAGACAGGATCAGGTCAGC								
NRAS exon3	FOR	CTTATTTAACCTTGGCAATAGCA								
NRAS exon3	REV + SEQ	GATTCAGAACACAAAGATCATCC								
HRAS exon2	FORW	GGCAGGAGACCCTGTAGGA								
HRAS exon2	REV + SEQ	AGCCCTATCCTGGCTGTGT								
HRAS exon3	FORW + SEQ	AGAGGCTGGCTGTGTGAACT								
HRAS exon3	REV	ATGCGCAGAGAGGACAGGA								
BRAF ex15	FORW	TGCTTGCTCTGATAGGAAAATG								
BRAF ex15	REV	AGCATCTCAGGGCCAAAAAT								
BRAF ex15	SEQ FOR	TGTTTTCCTTTACTTACTACACCTCA								
PIK3CA EX9	FORW	GGGAAAAATATGACAAAGAAAGC								
PIK3CA EX9	REV	CTGCTTTATTTATTCCAATAGGTATGG								
PIK3CA EX9	SEQ FOR	TAGCTAGAGACAATGAATTAAGGGAAA								
PIK3CA EX20	FORW	CTCAATGATGCTTGGCTCTG								
PIK3CA EX20	REV	TGGAATCCAGAGTGAGCTTTC								
PIK3CA EX20	SEQ FOR	TTGATGACATTGCATACATTCG								

Supplementary Data 3. Sanger sequencing primers's sequencees.

Supplementary Data 4. siRNA sequencees.								
For each target gene a pool of 4 siRNA was used. The table lists the								
sequencees of each siRNA included in the corresponding pool.								
siRNA Target	siRna Sequence (5'->3')							
ALK	GACAAGAUCCUGCAGAAUA							
ALK	GGAAGAGUCUGGCAGUUGA							
ALK	GCACGUGGCUCGGGACAUU							
ALK	GGUCAUAGCUCCUUGGAAU							
FGFR2	CCAAAUCUCUCAACCAGAA							
FGFR2	GAACAGUAUUCACCUAGUU							
FGFR2	GGCCAACACUGUCAAGUUU							
FGFR2	GUGAAGAUGUUGAAAGAUG							
KIT	AAACACGGCUUAAGCAAUU							
KIT	GAACAGAACCUUCACUGAU							
KIT	GGGAAGCCCUCAUGUCUGA							
KIT	GCAAUUCCAUUUAUGUGUU							
NTRK1	GGACAACCCUUUCGAGUUC							
NTRK1	CCAGUGACCUCAACAGGAA							
NTRK1	CCACAAUACUUCAGUGAUG							
NTRK1	GAAGAGUGGUCUCCGUUUC							
NTRK2	GAACAGAAGUAAUGAAAUC							
NTRK2	GUAAUGCUGUUUCUGCUUA							
NTRK2	GCAAGACACUCCAAGUUUG							
NTRK2	GAAAGUCUAUCACAUUAUC							
PDGFRA	CAUCAGAGCUGGAUCUAGA							
PDGFRA	GGCCUUACUUUAUUGGAUU							
PDGFRA	GAGCUUCACCUAUCAAGUU							
PDGFRA	CCUCUAGGAAUGACGGAUU							
RET	GCAAAGACCUGGAGAAGAU							
RET	GCACACGGCUGCAUGAGAA							
RET	UUAAAUGGAUGGCAAUUGA							
RET	GCGAGGAGAUGUACCGCCU							