IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

The clinical and biological significance of MIR-224 expression in colorectal cancer metastasis

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1615636 since 2017-01-27T16:21:38Z

Published version:

DOI:10.1136/gutjnl-2015-309372

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)

13 December 2025

IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

This is the author's final version of the contribution published as:

Ling, Hui; Pickard, Karen; Ivan, Cristina; Isella, Claudio; Ikuo, Mariko; Mitter, Richard; Spizzo, Riccardo; Bullock, Marc D.; Braicu, Cornelia; Pileczki, Valentina; Vincent, Kimberly; Pichler, Martin; Stiegelbauer, Verena; Hoefler, Gerald; Almeida, Maria I.; Hsiao, Annie; Zhang, Xinna; Primrose, John N.; Packham, Graham K.; Liu, Kevin; Bojja, Krishna; Gafà, Roberta; Xiao, Lianchun; Rossi, Simona; Song, Jian H.; Vannini, Ivan; Fanini, Francesca; Kopetz, Scott; Zweidler-McKay, Patrick; Wang, Xuemei; Ionescu, Calin; Irimie, Alexandru; Fabbri, Muller; Lanza, Giovanni; Hamilton, Stanley R.; Berindan-Neagoe, Ioana; Medico, Enzo; Mirnezami, Alex H.; Calin, George A.; Nicoloso, Milena S.. The clinical and biological significance of MIR-224 expression in colorectal cancer metastasis. GUT. 65 (6) pp: 977-989.
DOI: 10.1136/gutjnl-2015-309372

The publisher's version is available at:
https://syndication.highwire.org/content/doi/10.1136/gutjnl-2015-309372

When citing, please refer to the published version.

Link to this full text:
http://hdl.handle.net/2318/1615636

This full text was downloaded from iris - AperTO: https://iris.unito.it/

iris - AperTO

University of Turin's Institutional Research Information System and Open Access Institutional Repository

THE CLINICAL AND BIOLOGICAL SIGNIFICANCE OF MIR-224 EXPRESSION IN COLORECTAL CANCER

Hui Ling [a *], *Karen Pickard* [i *], *Cristina Ivan* [b *], *Claudio Isella* [l, m], *Mariko Ikuo* [a, v], *Richard Mitter* [j], *Riccardo Spizzo* [a, n], *Marc Bullock* [i], *Cornelia Braicu* [q], *Valentina Pileczki* [q], *Kimberly Vincent* [a], *Martin Pichler* [a, w], *Verena Stiegelbauer* [w], *Gerald Hoefler* [x], *Maria I. Almeida* [a], *Annie Hsiao* [a], *Xinna Zhang* [b], *John Primrose* [i, k], *Graham Packham* [i], *Kevin Liu* [a], *Krishna Bojja* [a], *Roberta Gafà* [o], *Lianchun Xiao* [c], *Simona Rossi* [a], *Jian H. Song* [d], *Ivan Vannini* [p], *Francesca Fanini* [p], *Scott Kopetz* [e], *Patrick Zweidler-McKay* [f], *Xuemei Wang* [c], *Calin Ionescu* [r, s], *Alexandru Irimie* [t], *Muller Fabbri* [h, p], *Giovanni Lanza* [o], *Stanley R. Hamilton* [g], *Ioana Berindan-Neagoe* [q, u], *Enzo Medico* [l, m], *Alex Mirnezami* [i, k †], *George A. Calin* [a, b †], *Milena S. Nicoloso* [a, n †]

a Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

b Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

c Division of Quantitative Science, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA

d Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

e Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

f Department of Pediatrics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

g Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

h Departments of Pediatrics, and Molecular Microbiology and Immunology, University of Southern California, Keck School of Medicine, The Saban Research Institute, Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA, USA

i Cancer Research UK Centre, University of Southampton Cancer Sciences Division, Somers Cancer Research Building, Southampton University Hospital NHS Trust, Tremona road, Southampton, SO16 6YD, UK

j Bioinformatics Unit, London Research Institute, Cancer Research UK, London, UK`

k Department of Surgery, Southampton University Hospital NHS Trust, Tremona road, Southampton, SO16 6YD, UK

l University of Torino, Department of Oncology, Torino, Italy

m IRCC, Institute for Cancer Research and Treatment, Candiolo, Torino, Italy

n Present address: Division of Experimental Oncology B, CRO, National Cancer Institute, Aviano, Italy

o Section of Pathology and Molecular Diagnostics, University of Ferrara, Ferrara, Italy

p Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) s.r.l., IRCCS, Gene Therapy Unit, Meldola (FC), Italy

q Department of Functional Genomics, The Oncology Institute, Cluj-Napoca, Romania.

r Surgical Clinic 1, Cluj County Hospital, Romania

s UMF Surgery Department 1, Cluj-Napoca, Romania

t Department of Surgical and Gynecology Oncology, University of Medicine and Pharmacy Iuliu Hatieganu, Cluj-Napoca, Romania

u Research Center for Functional Genomics, Biomedicine and Translational Medicine, University of Medicine and Pharmacy Iuliu Hatieganu, Cluj-Napoca, Romania

v Laboratory of Microbiology, Graduate School of Pharmaceutical Sciences, University of Tokyo, Tokyo, Japan

w Division of Oncology, Medical University of Graz, Austria

x Institute of Pathology, Medical University of Graz, Austria

* These authors contributed equally to this work

† Senior authors with equal contribution to this work

Correspondence to: **George A. Calin**, MD, PhD, Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA; email: gcalin@mdanderson.org; phone: +1-713-792-5461; **Milena S. Nicoloso**, MD, Division of Experimental Oncology B, CRO, National Cancer Institute, Aviano, Italy; email: mnicoloso@cro.it; phone: +39-0434-659820; **Alex Mirnezami**, MD, PhD, Cancer Research UK Centre, University of Southampton Cancer Sciences Division, Somers Cancer Research Building, Southampton University Hospital NHS Trust, Tremona road, Southampton, SO16 6YD, UK; email: A.H.Mirnezami@soton.ac.uk; phone: +39-0434-659820; A.H.Mirnezami@soton.ac.uk; +23-8079-5170.

Abstract

Background

MicroRNAs (miRNAs) are frequently altered in human cancers, and their expression levels can be used as prognostic markers. We aim to explore the clinical significance of miRNAs in colorectal cancer (CRC) metastasis.

Methods

We performed miRNA microarray in primary CRC tissues from patients with or without metastasis. The most differentially expressed miRNAs were validated in an independent set of 85 CRC samples by real time polymerase chain reaction (qRT-PCR), and tested for metastatic activity. We identified miRNA targets by prediction algorithms, qRT-PCR, western blot and luciferase assay. Clinical association was tested with six sets of CRC cases (n=449) including The Cancer Genome Atlas consortium. We used the Kaplan-Meier method and log-rank test to assess the difference in survival between patients with low or high levels of miR-224 and/or SMAD4 and CDH1 expression.

Results

We identified that miR-224 expression increases consistently with tumor burden and microsatellite stable (MSS) status, and miR-224 enhances CRC *in vitro* cell motility, and *in vivo* metastasis using an orthotopic CRC mouse model. We identified SMAD4 and CDH1 as miR-224 targets *in vitro*, but only SMAD4 showed negative correlation (Spearman Rs=-0.44, p<.0001) with miR-224 levels in patient samples. Patients with high miR-224 levels display shorter overall survival in multiple CRC cohorts (p=.0259, .0137, .0207, .0181, .0331 and .0037 respectively), and shorter metastasis-free survival (hazard ratio 6.51, 95% CI 1.97-21.51, p=.0008). In the TCGA set, combined analysis of miR-224 with SMAD4 expression improved the survival correlation (hazard ratio 4.12, 95% CI 1.1-15.41, p=.0175).

Conclusions

MiR-224 is specifically increased in MSS CRC, and acts to promote CRC metastasis through regulation of SMAD4. MiR-224 expression in primary CRC, alone or combined with its targets, may have prognostic value for CRC patient survival.

Keywords: miR-224; SMAD4; microsatellite stability; colorectal cancer; survival

Colorectal cancer (CRC) is the second leading cause of cancer death among adults, with over 1.6 million new cancer cases and 580,350 deaths estimated to have occurred in the United States in 2013 [1]. Metastatic spread of tumor cells remains the ultimate cause of cancer-related death in most CRC cases. While most cases of localized CRC (stage I-II) are curable by surgical excision, only about 70% of stage III CRC cases with regional lymph node metastasis are curable by surgery combined with adjuvant chemotherapy. Advanced metastatic disease (stage IV), despite improved survival due to recent advances in chemotherapy and targeted agents, remains largely incurable [2],[3]. Therefore, it is of critical importance to understand the key molecular switches involved in CRC metastasis, and identify biomarkers for CRC malignancies and prognostic markers for patient survival.

MicroRNAs (miRNAs) are 18- to 25-nucleotide RNAs that control gene expression at the post-transcriptional level. Based on sequence complementarity, miRNAs bind to targeted protein-coding genes, prevalently at their 3' untranslated region (UTR), and consequently affect mRNA stability or interfere with protein translation [4]. The functional importance of miRNAs in physiological and pathological conditions has been widely appreciated. MiRNAs are differentially expressed in normal and tumor tissues, and unique miRNA expression patterns have been characterized in many cancer types including CRC [5-7]. However, CRC metastasis-related miRNAs and their biological roles in CRC metastasis remain to be identified. In the invasive regions of primary CRC, organized structure of the tumor is lost: adhesion molecules that maintain cell–cell contact are downregulated, whereas molecules responsible for invasive and migratory behavior are upregulated [8]. These findings suggest that we could decipher tumor metastatic potential from the analysis of miRNA expression in primary CRC.

In this study, we aimed to identify miRNAs associated with CRC metastasis, and explore the biological significance as well as diagnostic and prognostic value of such miRNAs.

Methods

Clinical specimens

In total six independent CRC cohorts (Italy set 1, Italy set 2, UK set, Romania set, Austria set, and TCGA set) of 449 tumor samples (**Table 1**) and 172 non-neoplastic mucosal tissues were included in this study. Samples were obtained from University of Ferrara (Italy set 1, 85 tumor samples and 25 non-neoplastic mucosae), Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) s.r.l., IRCCS (Italy set 2, 68 tumor samples and 64 paired non-neoplastic mucosae), University of Southampton (UK set, 41 colon tumor samples), the Oncology Institute Cluj-Napoca (Romania set, 38 tumor samples and 23 paired non-neoplastic mucosae), and Medical University of Graz (Austria set, 74 tumor samples and 60 paired non-neoplastic mucosae) respectively. Samples from patients with biopsy proven CRC were obtained fresh at the time of surgery and snap frozen prior to being deposited. Tumor tissue, and uninvolved proximal mucosa were obtained. All CRC specimens were reviewed by pathologists, and contained more than 70% tumor cells as assessed by haematoxylin and eosin stain. Tumors used were adenocarcinomas only. Exclusion criteria included evidence of a hereditary tumour, presence of multiple or mucinous tumours, tumours with histologically identified extensive necrosis. Tumors were classified according to the World Health Organization pathologic classification system. Microsatellite analysis was evaluated with a fluorescence based PCR method using the five markers of the Bethesda panel (D5S346, D17S250, D2S123, BAT25 and BAT26) plus BAT40. According to the guidelines of the International Workshop of Bethesda [9], tumors were classified as MSI-H (instability at 2 or more loci), MSI-L (instability at single locus) or MSS (no instability). Patient's clinical information was registered and follow up data was recorded at checkup. All samples were obtained with patient's informed consent. The Institutional research and ethics committee approved this study.

TCGA dataset analysis

We obtained clinical information, and miRNA/mRNA expression from the published TCGA data (TCGA set, 143 CRC cases) [10]. Survival information was updated from TCGA Data Portal

(https://tcga-data.nci.nih.gov/tcga/). Patients that were recorded with 0 “days_to_last_follow_up”, and diagnosed after 82 years and died were excluded for homogeneity between cohorts. Level 3 Illumina RNASeq was used for miRNA, and mRNA expression.

Procedures (summarized in Figure 1)

Total RNA was isolated from snap frozen patient samples or from CRC cells using TRIzol reagent (Invitrogen) for Italy set 1, Italy set 2, and Romania set, and using RNAqueous-Micro Kit for UK set. Formalin-fixed paraffin embedded samples were use for RNA extraction in Austria set. RNA quantification was performed by Nanodrop, and quality was checked by agarose gel electrophoresis or a 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). We used a custom miRNA microarray containing quadruplicates of 389 human miRNA probes for miRNA profiling as previously described [11].

Based on the microarray results, we further examined the metastasis-related miRNAs using real time polymerase chain reaction (qRT-PCR) analysis with TaqMan® MiRNA assays (Life Technologies, Grand Island, NY, USA).

Biological functions of miRNAs were tested using *in vitro* motility assays, and an orthotopic mouse model of CRC. Transiently and stably transfected CRC cell lines were established for the functional and mechanistic studies. To identify the mRNA targets, we used a prediction algorithm, and tumor metastasis PCR array for the screening, and qRT-PCR, western blot, luciferase assay, immunocytochemistry, and rescue experiments for validation. Clinical correlation was tested with multiple sets of CRC samples, and publically available TCGA data. Other methods are shown in **Supplementary methods** (available online), and primers are listed in **Supplementary Table 1**.

***In vivo* study**

Six to eight weeks old SCID mice (Charles River, UK) were anaesthetized prior to a midline laparotomy and exteriorization of the caecum. A 1:1 suspension of cells and Matrigel was injected submucosally into the caecal wall under magnified vision, raising a bleb on the caecum. For each

animal, $5x10^6$ cells stably overexpressing GFP tagged miR-224 or control miRNA were implanted orthotopically, with the entire experiment conducted in duplicate. Primary tumors grew in all animals. When showing signs of disease or more than 10% weight loss, mice were humanely culled, and colon, liver and lungs were harvested. Excised tissue was paraffin embedded, and stained with haematoxylin and eosin. Pulmonary metastases were manually counted in 10 random fields, while liver replacement was quantified using the program ImageJ. Mean tumor burden was calculated, and results presented relative to controls.

Statistical analysis

Differences between groups were analyzed using Student t test (2-tailed), assuming equal or unequal variance determined by the F-test of equality of variances. Graphics represent the mean ± standard deviation, unless otherwise stated. The Spearman correlation coefficients were computed to assess the correlation between expression level of miR-224 and its target genes in clinical samples. For survival analysis, we divided patients into low/high groups using as cut-off the value that optimally separated the patients, and used the Kaplan-Meier method to estimate the survival curves, and the log-rank test for the comparison. A p value <.05 was considered statistically significant.

Results

Identification of metastasis-related miRNAs in colorectal cancer

To identify microRNAs related to CRC metastasis, we performed miRNA microarray using primary CRCs from patients (Italy set 1) with (n=4) or without (n=8) metastasis at diagnosis, and in commercially available cell lines derived from a primary CRC lesion (SW480) or its metastatic dissemination to the lymph node (SW620). According to the microarray data (**Supplementary Figure 1A and Supplementary Table 2**) 11 miRNAs significantly correlated with the presence of metastasis. We decided to validate them by qRT-PCR in a larger cohort of samples (Italy set 1, comprising 85 primary CRC tumors and 25 matched adjacent non-neoplastic colon mucosae). Four of the tested miRNAs (miR-141, miR-181b, miR-221, and miR-224; **Table 2**) showed a significant increase in expression levels in primary CRC at advanced stages, including metastatic dissemination to lymph nodes and to distant organs (stage III-IV) compared to early stages (stage I-II), as well as in the CRC metastasis-derived SW620 cell line compared to primary CRC-derived SW480 cell line. Among these, miR-181b, miR-221 and miR-224 were significantly increased in neoplastic CRC tissue compared to normal mucosa. Higher levels of miR-224 in advanced stages, and in tumor versus normal tissue were supported by analysis from multiple cohorts of primary CRC, an independent dataset from The Cancer Genome Atlas (TCGA) consortium, and cell lines with different metastatic features (**Table 1, Figure 2A, Supplementary Figure 1B,** and **Supplementary Figure 2**).

Association of miR-224 expression with microsatellite status and tumor site

Unlike miR-181b and miR-221, miR-224 levels were significantly higher in microsatellite stable (MSS) samples compared to microsatellite instability-high (MSI-H) tumors, concordant with the notion that MSI-H CRCs are less aggressive and less prone to metastatic spread than MSS tumors [12] (**Figure 2B**). Also consistent with the reported association of CRC localization with microsatellite status [13], miR-224 levels were significantly lower in right colon tumors than those occurring in the left colon and rectum (**Figure 2B**). These findings were validated in TCGA dataset comprising 143 samples (**Figure 2C**), and Italy set 2 comprising 67 samples (**Figure 2D**). Concomitantly, *in situ* hybridization showed

stronger (fold change = 3.6) miR-224 staining in epithelial cells of MSS CRCs compared with normal and MSI-H tumor samples (**Figure 2E** and **Supplementary Figure 3**).

Pro-metastatic activity of miR-224 *in vitro* and *in vivo*

We selected five miRNAs (the four miRs from primary screening, miR-141, miR-181b, miR-221, and miR-224, as well as miR-222 that contains identical seed sequences with miR-221) for functional test using transwell-based assays (migration, haptotaxis, and invasion) with HCT116 cells. Cell motility increased consistently only upon overexpression of miR-224 and miR-141 by either transient or stable transfection (**Figure 3A, Supplementary Figure 4,** and **Supplementary Figure 5A**). Next, using an orthotopic CRC SCID mouse model, we investigated *in vivo* activity of miR-224. Direct caecal implantation of HCT116 cells stably over-expressing miR-224 (n=4) resulted, at 5 weeks post-procedure, in a greater number and size of metastatic tumor deposits in the liver and lungs compared to control cells. Liver replacement was almost complete in some mice implanted with HCT116-miR-224, suggesting a rapid metastatic and growth process. Representative macroscopic and H&E staining of tumor deposits are shown in **Figures 3B** and **3C**. Similarly, stable ectopic expression of miR-224 (75-fold induction) with a GFP co-expression plasmid construct in RKO, a CRC cell line with low miR-224 expression (**Supplementary Figure 1B**) promoted both *in vitro* cell motility (**Supplementary Figure 5B)** and *in vivo* tumor metastasis to the liver in the orthotopic SCID mouse model (**Figure 3D)**. Although RKO lung metastases were not as overt as those with HCT116 cells, by immunostaining with anti-GFP antibodies, we observed a significantly greater number of miR-224-RKO cells disseminated to the lung compared to control RKO cells (**Figure 3E**). The increased metastatic phonotype observed using two different cell models strongly supports the pro-metastatic function of miR-224 in CRC.

SMAD4 and CDH1 as miR-224 targets

We performed a PCR array comprising 87 metastasis-related genes on SW480 cells. Among the 13 genes that were reduced more than half by miR-224 overexpression (**Supplementary Table 3**), only

CDH1 and SMAD4 were predicted miR-224 targets by the miRGEN database (http://diana.pcbi.upenn.edu/cgi-bin/miRGen/v3/Targets.cgi) (**Figure 4A**). We validated these findings in multiple CRC cell lines. Overexpression of miR-224 decreased the mRNA (**Supplementary Figure 6)** and protein (**Figure 4B**) expression of CDH1 and SMAD4, both upon transient and stable miRNA transduction. Inversely, knockdown of miR-224 in HCT116 and AAC1/82 caused an upregulation of SMAD4 protein expression (**Figure 4C**). Next, we generated luciferase constructs containing SMAD4 and CDH1 3' UTR sequences at the 3' end of luciferase gene (pGL3-CDH1, pGL3-SMAD4 constructs A and B, and their respective miR-224 target mutants) [14]. Co-transfection with synthetic miR-224 precursor reduced the luciferase activity of pGL3-CDH1, but not the mutant construct with the predicted interaction sequence deleted (**Figure 4D**). The reporter activity of pGL3-SMAD4 construct A, which was reduced to half by miR-224 overexpression, reverted after deletion of binding site 1 but not of binding site 2 (**Figure 4D**). On the contrary, pGL3-SMAD4 construct B, containing binding site 3 was not affected by miR-224 (**Figure 4D**). Independent experiments with point mutants of the binding sites consistently showed that binding site 1 is the interaction site for miR-224 action (**Supplementary Figure 7**). As a further proof of SMAD4 regulation by miR-224, in the HCT116 cells transfected with GFP-miR construct, we observed a mutually exclusive expression pattern of GFP (indicating miR-224 expression) and SMAD4 (**Figure 4E**).

Inverse correlation of miR-224 and SMAD4 in clinical samples

We moved on to measure SMAD4 and CDH1 protein levels in stage I-II CRC samples with low miR-224 levels and in stage IV samples with high miR-224 levels selected from the Italian CRC set 1. Consistent with *in vitro* findings, we detected an inverse correlation between miR-224 levels and SMAD4 protein expression (**Figure 5A**). To exclude the possibility that SMAD4 expression arises from tissues other than the epithelial cells, we performed laser capture microdissection (LCM) in normal colon and CRC tissues, and observed a reciprocal expression of miR-224 and SMAD4 mRNA (**Figure 5B**). However, we did not find an inverse correlation between CDH1 and miR-224 (**data not shown**). Consistently, also in the TCGA consortium data (n=133) we found an inverse correlation

between miR-224 and SMAD4 (Spearman Rs=-0.44, p<.0001), but not with CDH1 (Spearman Rs=0.32, p=.0002) (**Supplementary Figure 8**).

SMAD4 as effector of miR-224 in promoting cell motility

We next investigated whether SMAD4 is the mediator of miR-224's pro-metastatic effect. Similar to what occurred after miR-224 transfection, silencing of SMAD4 by RNA interference increased HCT116 cell motility (**Figure 5C**). Moreover, exogenous SMAD4 expression with a miR-224-resistant SMAD4 construct (SMAD4 without its 3'UTR containing the putative miR-224 target site) abrogated miR-224's ability to promote cell migration and invasion (**Figure 5D**). Taken together, the phenocopy and the rescue experiment support the hypothesis that SMAD4 is a key effector of miR-224's pro-metastatic capacity.

Association of miR-224 with CRC patient survival

To determine the clinical significance of such findings, we performed patient survival analysis in five CRC cohorts with available follow-up information (**Table 1**). In all sets of CRC samples, patients with high miR-224 expression had shorter overall survival compared to those with low miR-224 levels with p=.0259 in TCGA dataset (n=143), p=.0137 in Italy set 1 (n=54), p=.0207 in Italy set 2 (n=68), p=.0181 in UK set (n=41), and the p=.0331 in Romania set (n=38) (**Figure 6A-6E and Table 3**). Multivariate analysis showed a trend of correlation but not reaching statistical significance, possibly due to the smaller size in the groups after stage separation. Therefore, to examine if miR-224 association with survival depends on stage, we used an Austrian cohort of samples comprising 74 colon tumors, with majority of them in stage III and IV. In the multivariate analysis, miR-224 showed prognostic value on overall survival (HR 2.36, 95% CI 1.32-4.21; p=.0037) independent of tumor stage (**Figure 6F** and **Table 3**). In addition, in the TCGA dataset, combined analysis of miR-224 with SMAD4 expression increased the separation of the survival curves by either gene alone in the TCGA dataset, and patients with miR-224 (high)/SMAD4 (low) had shorter survival compared to those with miR-224 (low)/SMAD4 (high) (HR 4.12, 95% CI 1.1-15.41; p=.0175) (**Figure 6G and Supplementary**

Figure 9). Interestingly, although we did not an observe inverse correlation for miR-224 and CDH1 expression, and CDH1 alone did not predict patient survival, combined analysis with CDH1 greatly improved the prediction power of miR-224 for patient overall survival (p=.0009) (**Supplementary Figure 10**). Furthermore, in the UK set (where complete clinical information was available), patients with high miR-224 expression in primary CRC had shorter metastasis-free survival than those with low miR-224 expression (HR 6.51, 95% CI 1.97-21.51; p=.0008) (**Figure 6H**). Notably, miR-224 showed higher sensitivity and specificity (Area Under Curve=.739) for metastasis-free survival than for overall survival in the receiver-operating characteristic (ROC) analysis (**Supplementary Figure 11**), further supporting a specific involvement of miR-224 in CRC metastasis.

Discussion

Reports on miR-224 involvement in CRC have been controversial. For instance, it was reported that miR-224 promotes CRC tumor growth in mice by repressing PHLPP1 and PHLPP2 [15], as well as that it suppresses tumor growth and metastasis *in vivo* [16]. In addition, one study found lower miR-224 expression in methotrexate-resistant colon cancer cells [17], while another study showed that a miR-224 mimic decreased the chemoradiosensitivity of CRCs [18]. The latter study also showed lower miR-224 expression in metastatic CRC cell lines versus non-metastatic cell lines, and in metastatic tumors in the lung versus primary CRC tumors. Although this discrepancy can be explained by the fact that in our approach we analyzed miR-224 expression in primary CRC samples, our *in vitro* and *in vivo* studies strongly support a pro-metastatic function of miR-224 rather than an anti-metastatic one. This pro-metastatic effect is also supported by our data from multiple CRC cohorts showing higher miR-224 expression in advanced stages, and shorter survival in patients with high miR-224 expression levels, consistent with the previous report showing that miR-224 expression increases with higher T stage, and in CRCs with distant metastasis [15].

Interestingly, among all the clinical parameters that we have analyzed (including tumor stage and tumor site), microsatellite status is the most significant factor associated with miR-224 levels in the 85 Italian CRC samples and the 143 TCGA cases. MSS CRCs are characterized by their unstable

chromosomal status, which has been linked with malignant phenotypes such as metastasis and suggested as an indicator of poor prognosis [19]. There is no report on such findings except one showing that miR-224 expression in CRCs with proficient DNA mismatch repair is more than twice that of CRC with defective DNA mismatch repair [20]. MSS CRCs account for 80% of all CRC cases, however a direct diagnosis for MSS CRC is still lacking. Current strategies characterize microsatellite status based on exclusion of multiple microsatellite instability marks. The specific involvement of miR-224 in CRC has the potential of developing into a reliable diagnostic marker for MSS CRCs.

SMAD4 protein is a transcription factor required for synergistic transcriptional activity in response to TGF-beta, and loss of SMAD4 in CRC was shown to switch TGF-beta function from tumor suppressor to a promoter of tumorigenicity and metastasis [21]. Loss of SMAD4 function due to either 18q deletion, reduced expression or protein inactivation has been associated with advanced and metastatic colorectal cancers [22, 23]. CRC tumors with or without 18q21 allelic imbalance showed no difference in SMAD4 levels, suggesting additional mechanisms of SMAD4 regulation [24]. Furthermore, while 18q21 allelic imbalance and SMAD4 mutations did not perform well as prognostic markers, the level of SMAD4 was found to have prognostic value for these CRC patients [24]. SMAD4 has been suggested as a miR-224 target in cancer and other disease models [25-27]. Recently, a report showed that miR-224 also targets SMAD4 in CRC models [28]. We provide solid evidence that SMAD4 is a miR-224 target by using target prediction, overexpression and knockdown cell models, luciferase assay, and expression correlation in clinical samples. Furthermore, we have validated SMAD4 as a miR-224 mediator for its pro-metastatic activity by phenocopy and rescue experiments. Interestingly, we also identified CDH1 as a miR-224 target, by target prediction, qRT-PCR, Western blot, and luciferase assay; however, CDH1 mRNA shows positive, instead of inverse, correlation with miR-224 levels in clinical samples. Despite the discrepancies, high miR-224 and low CDH1 show a strikingly better prediction for CRC patient overall survival than either gene alone, suggesting that miR-224-CDH1 regulation also plays a role in CRC malignancies.

We are aware of the limitations of the present study. First, although we used multiple independent sets of CRC samples, larger cohorts are needed to validate findings on miR-224

expression and CRC patient survival, and on the prognostic advantage of combining the expression levels of miR-224 and its mRNA targets such as SMAD4 and CDH1. Second, the survival analysis in this study could be confounded by treatment received by patients, and the therapeutic interference should be considered when performing larger scale of retrospective or prospective biomarker studies.

Nevertheless, this study has several clinical implications. First, the specific involvement of miR-224 in MSS CRCs indicates its potential to be developed into a diagnostic marker for these patients. Secondly, miR-224 alone or in combination with its target genes (SMAD4, CDH1, or possibly other targets) may serve as prognostic marker for CRC patients. The recent finding that CRC patients with MSS status and loss of SMAD4 expression had significantly worse survival [29] adds further support to the prognostic values of miR-224 and SMAD4. Thirdly, high miR-224 expression in advanced CRCs and its pro-metastatic effects suggest that this miRNA could be an ideal candidate target for CRC treatment, and that this strategy would be most applicable to patients whose tumors have low SMAD4 expression. The recent success of miravirsen (an LNA anti-miR-122) in a clinical trial for treating HCV infection [30] indicates technical plausibility of anti-miR-224 treatment.

Conflicts of interest

We declare that we have no conflicts of interest.

Funding

Dr Calin is The Alan M. Gewirtz Leukemia & Lymphoma Society Scholar. Work in Dr. Calin's laboratory is supported in part by the NIH/NCI grants 1UH2TR00943-01 and 1 R01 CA182905-01, Developmental Research Awards in Prostate Cancer, Multiple Myeloma, Leukemia (P50 CA100632), Head and Neck (P50 CA097007) and Brain (2P50CA127001) SPOREs, a SINF MDACC_DKFZ grant in CLL, a SINF grant in colon cancer, and a Kidney Cancer Pilot Project. The Center for Cancer Epigenetics Pilot project, a 2014 Knowledge GAP MDACC grant, the Laura and John Arnold Foundation, the RGK Foundation and the Estate of C. G. Johnson, Jr and by the CLL Global Research Foundation. Dr. Ling is an Odyssey Fellow, and funding as an Odyssey Fellow was supported by the Odyssey Program and The Estate of C. G. Johnson, Jr. at The University of Texas MD Anderson Cancer Center. Dr. Pichler is supported by an Erwin-Schroedinger Scholarship of the Austrian Science Funds (project no. J3389-B23). Dr. Almeida is supported as a fellow by the Foundation for Science and Technology, Portugal (SFRH/BPD/91011/2012). Dr. Hamilton is the recipient of the Frederick F. Becker Distinguished University Chair in Cancer Research from The University of Texas. Work in Dr. Fabbri's lab is supported in part by award number P30CA014089 from the National Cancer Institute. Work in Dr. Medico's laboratory is supported in part by funds from AIRC (IG n. 12944 and 5x1000 project n. 9970), FPRC and Regione Piemonte (E-LAB). Dr. Isella is a University of Torino Fellow, supported in part by Regione Piemonte (E-LAB). Dr. Pickard and Dr. Mirnezami are supported by grant funding from Wessex Medical Research and Cancer Research UK (C28503/A10013). Work in Dr Nicoloso's lab is supported in part by funds from AIRC (MFAG n. 13589), 5x1000 to CRO and Marie Curie Actions (CIG n. 303877).

Notes

H. Ling, K. Pickard, and C. Ivan contributed equally to this work.

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report or the decision to submit this manuscript for publication.

The author contributions are as follows: AM, GAC, and MSN designed the study; MSN, HL, KP, MI, RS, MB, CB, VP, MIA, KV, VX, AH, XZ, KL, JHS, SK, PZ-M, and IB-N performed the wet-lab experiments; JP, GP, RG, IV, MP, GH, FF, MF, AI, C. Ionescu, GL, SRH, IB-N, and AM obtained samples and clinical data; C. Ivan, RM, LX, HL, and XW did statistical analysis; C. Ivan, C. Isella and EM performed TCGA analysis; HL, KP, GAC, AM and MSN did data analysis and interpretation; HL, KP, MI and MSN write the initial draft; GAC, MSN, AM, KV, C. Isella and EM revised the report. All authors reviewed this report and approved the final version.

Affiliations of authors

Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (HL, MI, RS, MIA, KV, MP, KB, AH, KL, SR, GAC, MSN); Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (C. Ivan, XZ, GAC); Division of Quantitative Science, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA (LX, XW); Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (JHS); Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (SK); Department of Pediatrics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (P Z-M); Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (SRH); Departments of Pediatrics, and Molecular Microbiology and Immunology, University of Southern California, Keck School of Medicine, The Saban Research Institute, Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA, USA (MF); Cancer Research UK Centre, University of Southampton Cancer Sciences Division, Somers Cancer Research Building, Southampton University Hospital NHS

Trust, Tremona road, Southampton, SO16 6YD, UK (KP, MB, JP, GP, AM); Bioinformatics Unit, London Research Institute, Cancer Research UK, London, UK (RM); Department of Surgery, Southampton University Hospital NHS Trust, Tremona road, Southampton, SO16 6YD, UK (JP, AM); University of Torino, Department of Oncology, Torino, Italy (C. Isella, EM); IRCC, Institute for Cancer Research and Treatment, Candiolo, Torino, Italy (C. Isella, EM); Division of Experimental Oncology B, CRO, National Cancer Institute, Aviano, Italy (RS, MSN); Section of Pathology and Molecular Diagnostics, University of Ferrara, Ferrara, Italy (RG, GL); Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) s.r.l., IRCCS, Gene Therapy Unit, Meldola (FC), Italy (IV, FF, MF); Department of Functional Genomics, The Oncology Institute, Cluj-Napoca, Romania (CB, VP, I B-N); Surgical Clinic 1, Cluj County Hospital, Romania (C. Ionescu); UMF Surgery Department 1, Cluj-Napoca, Romania (C. Ionescu); Department of Surgical and Gynecology Oncology, University of Medicine and Pharmacy Iuliu Hatieganu, Cluj-Napoca, Romania (AI); Research Center for Functional Genomics, Biomedicine and Translational Medicine, University of Medicine and Pharmacy Iuliu Hatieganu, Cluj-Napoca, Romania (I B-N); Laboratory of Microbiology, Graduate School of Pharmaceutical Sciences, University of Tokyo, Tokyo, Japan (MI); Division of Oncology, Medical University of Graz, Austria (MP, VX); Institute of Pathology, Medical University of Graz, Austria (GH).

Figure legends

Figure 1. A schematic description of the workflow in this study.

Figure 2. miR-224 expression in primary CRC samples. (A) Higher miR-224 expression in advanced stages from multiple CRC cohorts, as determined by Taqman qRT-PCR or obtained from TCGA dataset. **(B)** Differential miR-224 expression in the Italian set 1 of CRC samples subdivided by MSS/MSI status or tumor site. **(C, D)** Validation of miR-224 association with MSS/MSI status and tumor site with TCGA dataset and Italian set 2. **(E)** *In situ* hybridization of the miR-224 in normal mucosae (N 1-4), MSI-H (MSI 1-4) and MSS CRC samples (MSS 1-4). MiR-224 expression was normalized by U6. Data in **A-D** are presented as box-whisker plots showing the five statistics (lower whisker is the 10^{th} percentile, lower box part is the 25^{th} percentile, solid line in box is the median, upper box part is 75th percentile, and upper whisker is 90^{th} percentile).

Figure 3. MiR-224 promotes CRC metastasis. (A) Transwell-based motility assays in HCT116 cells transiently transfected with miRNA mimics. Motility is expressed as fold change compared to control cells. Data are presented as mean ± SD of three independent experiments each in triplicates. A representative image is shown under the bar chart for each treatment condition. **(B-D)** Ectopic miR-224 expression promotes CRC metastasis to the liver and lung in both HCT116 (**B** and **C,** n=4 in each group) and RKO (**D** and **E,** n=6 in each group) cell systems. Orthotopic SCID mouse models were established with CRC cells stably transfected with miR-224 or control (0.5 million cells). After sacrificing mice, liver and lungs were dissected and subjected to histological evaluation with Haematoxylin and Eosin. The number of pulmonary deposits and degree of liver replacement was calculated using imageJ analysis software. Results are presented as relative to untransfected cells. *, $p<.05$. **(E)** Lung immunostaining with antibody against GFP, which is coexpressed by miR ctrl or miR-224 constructs, to detect CRC cell infiltrates.

Figure 4. MiR-224 regulates SMAD4 and CDH1 expression. (A) MiR-224 interaction sites with CDH1 3'UTR according to microRNA.org (left panel), and with SMAD4 3'UTR according to TargetScan (righ panel). Two SMAD4 constructs were produced as indicated (construct A, containing sites 1 and 2, and construct B, containing site 3). **(B)** MiR-224 reduces SMAD4 and CDH1 protein expression in SW480 and HCT116 cells after transient transfection, as well as in HCT116 stably transfected cells. SMAD4 was never detected at the protein level in SW480 cells. **(C)** AntimiR-224 treatment increases the protein expression levels of SMAD4 in HCT116 and AAC1/82 cells. Western blot bands were quantified by image J and the number shows the relative expression. **(D)** Luciferase activity of pGL3-CDH1 and SMAD4 constructs containing miR-224 predicted binding sites (see **A** for annotation) after scrambled or miR-224 transfection in HCT116 cells. Luciferase activity is expressed as fold change relative to scrambled. One representative experiment is shown (presented as mean ± SD) out of at least three independent experiments performed in quadruplicates. *, $p<.05$. **(E)** Immunocytochemistry for GFP-miR-224 (green) and SMAD4 (red) in stably transfected HCT116 cells.

Figure 5. Inverse correlation of miR-224 and SMAD4 in clinical samples, and rescue experiment. (A) SMAD4 protein expression in Italian cohort 1 CRC samples with different stages. **(B)** Inverse expression profile of miR-224 and SMAD4 protein in epithelial cells of normal tissues and CRC samples by LCM. **(C)** SMAD4 siRNA enhanced CRC cell motility, similar with that observed in miR-224 overexpressing experiments. **(D)** Enforced SMAD4 expression abrogated miR-224's effect on promoting HCT116 cell migration and invasion. The experiments were performed twice in triplicates. One representative experiment is shown as mean ± SD.

Figure 6. MiR-224 and its target gene SMAD4 as prognostic markers in CRC patients. (A-F) Association of miR-224 expression in primary CRC with overall survival in multiple CRC cohorts. **(G)** Combined expression of miR-224 and SMAD4 improved the separation curve of patient overall survival. **(H)** Metastasis-free survival analysis of miR-224 association in the UK set. A statistically

determined optimal cut-off was used to stratify the patient groups, and the log-rank test was used for survival analysis.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013;63(1):11-30.

2. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. N Engl J Med 2009;361(25):2449-60.

3. Lindmark G, Gerdin B, Pahlman L*, et al.* Prognostic predictors in colorectal cancer. Dis Colon Rectum 1994;37(12):1219-27.

4. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5(7):522-31.

5. Slaby O, Svoboda M, Michalek J*, et al.* MicroRNAs in colorectal cancer: translation of molecular biology into clinical application. Mol Cancer 2009;8:102.

6. Bandres E, Cubedo E, Agirre X*, et al.* Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Mol Cancer 2006;5:29.

7. Ma Y, Zhang P, Yang J*, et al.* Candidate microRNA biomarkers in human colorectal cancer: systematic review profiling studies and experimental validation. Int J Cancer 2012;130(9):2077-87.

8. Brabletz T, Hlubek F, Spaderna S*, et al.* Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. Cells Tissues Organs 2005;179(1-2):56-65.

9. Boland CR, Thibodeau SN, Hamilton SR*, et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58(22):5248-57.

10. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487(7407):330-7.

11. Liu CG, Calin GA, Meloon B*, et al.* An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci U S A 2004;101(26):9740-4.

12. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005;23(3):609-18.

13. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993;260(5109):816-9.

14. Cimmino A, Calin GA, Fabbri M*, et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102(39):13944-9.

15. Liao WT, Li TT, Wang ZG*, et al.* microRNA-224 Promotes Cell Proliferation and Tumor Growth in Human Colorectal Cancer by Repressing PHLPP1 and PHLPP2. Clin Cancer Res 2013;19(17):4662-72.

16. Yuan K, Xie K, Fox J*, et al.* Decreased Levels of miR-224 and the Passenger Strand of miR-221 Increase MBD2, Suppressing Maspin and Promoting Colorectal Tumor Growth and Metastasis in Mice. Gastroenterology 2013;145(4):853-864 e9.

17. Mencia N, Selga E, Noe V*, et al.* Underexpression of miR-224 in methotrexate resistant human colon cancer cells. Biochem Pharmacol 2011;82(11):1572-82.

18. Salendo J, Spitzner M, Kramer F*, et al.* Identification of a microRNA expression signature for chemoradiosensitivity of colorectal cancer cells, involving miRNAs-320a, -224, -132 and let7g. Radiother Oncol 2013; 10.1016/j.radonc.2013.06.032.

19. Hemminki A, Mecklin JP, Jarvinen H*, et al.* Microsatellite instability is a favorable prognostic indicator in patients with colorectal cancer receiving chemotherapy. Gastroenterology 2000;119(4):921-8.

20. Oberg AL, French AJ, Sarver AL*, et al.* miRNA expression in colon polyps provides evidence for a multihit model of colon cancer. PLoS One 2011;6(6):e20465.

21. Zhang B, Halder SK, Kashikar ND*, et al.* Antimetastatic role of Smad4 signaling in colorectal cancer. Gastroenterology 2010;138(3):969-80 e1-3.

22. Miyaki M, Iijima T, Konishi M*, et al.* Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. Oncogene 1999;18(20):3098-103.

23. Xie W, Rimm DL, Lin Y*, et al.* Loss of Smad signaling in human colorectal cancer is associated with advanced disease and poor prognosis. Cancer J 2003;9(4):302-12.

24. Alazzouzi H, Alhopuro P, Salovaara R*, et al.* SMAD4 as a prognostic marker in colorectal cancer. Clin Cancer Res 2005;11(7):2606-11.

25. Wang Y, Ren J, Gao Y*, et al.* MicroRNA-224 targets SMAD family member 4 to promote cell proliferation and negatively influence patient survival. PLoS One 2013;8(7):e68744.

26. Yao G, Yin M, Lian J*, et al.* MicroRNA-224 is involved in transforming growth factor-beta-mediated mouse granulosa cell proliferation and granulosa cell function by targeting Smad4. Mol Endocrinol 2010;24(3):540-51.

27. Lichner Z, Mejia-Guerrero S, Ignacak M*, et al.* Pleiotropic action of renal cell carcinoma-dysregulated miRNAs on hypoxia-related signaling pathways. Am J Pathol 2012;180(4):1675-87.

28. Zhang GJ, Zhou H, Xiao HX*, et al.* Up-regulation of miR-224 promotes cancer cell proliferation and invasion and predicts relapse of colorectal cancer. Cancer Cell Int 2013;13(1):104.

29. Roth AD, Delorenzi M, Tejpar S*, et al.* Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. J Natl Cancer Inst 2012;104(21):1635-46.

30. Janssen HL, Reesink HW, Lawitz EJ*, et al.* Treatment of HCV infection by targeting microRNA. N Engl J Med 2013;368(18):1685-94.

Table 1

	Italy set 1 (n=85)	Italy set 2 (n=68)	UK set (n=41)	Romania set (n=38)	Austria set (n=74)	TCGA set (n=143)
Role of cohort in present study	Used for normal v. tumor (25 normal; 85 tumor) Used for stage association (n=85) Used for site association (n=85) Used for microsatellite status association (n=85) Used for survival analysis (n=54)	Used for normal v. tumor (n=64, paired) Used for stage association (n=68) Used for site association (n=67) Used for survival analysis (n=68)	Used for stage association (n=41) Used for survival analysis (n=41)	Used for normal v. tumor (n=23, paired) Used for stage association (n=38) Used for survival analysis (n=38)	Used for normal v. tumor (n=60, paired) Used for survival analysis (n=74)	Used for stage association (n=141) Used for site association (n=136) Used for microsatellite status association (n=143) Used for survival analysis (n=143)
Median age (range) years	73 (42-93)	67(41-88)	70 (38-94)	69.5 (38-81)	63 (38-80)	68 (34-89)
Sex						
Male	40 (47%)	40 (59%)	20 (49%)	25 (66%)	46(62%)	75 (52%)
Female	45 (53%)	28 (41%)	21 (51%)	13 (34%)	28 (38%)	68 (48%)
Tumor location						
Right	51 (60%)	20 (29%)	14 (34%)	10 (26%)	27 (36%)	58 (40%)
Left	23 (27%)	27 (40%)	27 (66%)	13 (34%)	47 (64%)	37 (26%)
Rectum	11 (13%)	20 (29%)		9 (24%)		43 (30%)
Not known		1 (2%)		6 (16%)		5 (3%)
Tumor stage						
I	3 (4%)	17 (25%)	4 (10%)	5 (13%)	0 (0%)	31 (22%)
II	33 (39%)	24 (35%)	18 (44%)	14 (37%)	7 (9%)	47 (33%)
III	20 (23%)	15 (22%)	19 (46%)	13 (34%)	23 (31%)	38 (27%)
IV	29 (34%)	12 (18%)		6 (16%)	44 (59%)	25 (17%)
not known						2 (1%)
Microsatellite status						
MSI-H	37 (44%)					21 (15%)
MSI-L	0 (0%)					19 (13%)
MSS	48 (56%)					103 (72%)

All data are n (%) unless stated otherwise

Table 1: Main characteristics of patients and tissue samples

Table 2

Table 2. Differentially expressed miRNAs in primary CRCs from patients or CRC cell lines

	Tum/Norm	p-value	**MSS/MSI**	p-value	**III-IV/I-II**	p-value	**SW620/SW480**
miR-141	**0.74**	**0.0000**	1.18	0.1411	**1.36**	**0.0065**	4.90
miR-181b	**1.28**	**0.0015**	0.93	0.5580	**1.44**	**0.0025**	4.51
miR-191	0.91	0.1283	1.22	0.0931	**1.28**	**0.0338**	0.52
miR-200a	**0.80**	**0.0370**	0.92	0.5013	1.25	0.0656	NOT AVAILABLE
miR-200b	**0.81**	**0.0019**	1.09	0.4497	1.23	0.0835	25.76
miR-203	**1.57**	**0.0020**	1.20	0.4218	**1.92**	**0.0024**	0.19
miR-215	**0.27**	**0.0000**	0.74	0.3673	1.15	0.6852	50.20
miR-221	**3.59**	**0.0000**	0.82	0.3669	**1.73**	**0.0057**	4.04
miR-222	**1.45**	**0.0044**	**0.62**	**0.0215**	**1.80**	**0.0023**	NOT AVAILABLE
miR-224	**3.76**	**0.0000**	**2.99**	**0.0000**	**1.54**	**0.0404**	1.49
miR-425	1.01	0.9484	0.98	0.8111	1.21	0.0863	NOT AVAILABLE

Red Bold: ratio larger than 1, and difference significant ($p < 0.05$)
Green Bold: ratio lower than 1, and difference significant ($p < 0.05$)

Table 3

		UNIVARIATE		MULTIVARIATE	
Cohort	**Variable**	**HR (95%CI)**	**p-value (log-rank)**	**HR (95%CI)**	**p-value(wald)**
TCGA	Age (<median vs >median)	1.02(0.37,2.83)	0.9614		
	Gender (Male vs Female)	1.32(0.49,3.55)	0.579		
	Tumor location (Left vs Right)	0.92(0.32,2.68)	0.8846		
	Tumor stage(III-IV vs I-II)	7.87(1.77,35.07)	0.0014 *	8.37(1.82,38.51)	0.0063 *
	Microsatellite status (MSS vs MSI)	2.896(0.66,12.75)	0.1597		
	miR-224(High vs Low)	2.99(1.08,8.23)	0.0259 *	2.88(0.97,8.56)	0.0571
Italian set 1	Age (<median vs >median)	0.81(0.39,1.71)	0.5817		
	Gender (Male vs Female)	1.56(0.72,3.39)	0.2598		
	Tumor location (Left vs Right)	0.74(0.35,1.57)	0.4284		
	Tumor Stage(Stage III-IV vs Stage I-II)	2.85(1.28,6.35)	0.0073	2.44(1.09, 5.48)	0.0302 *
	MSI_status (MSS vs MSI)	2.04(0.87,4.81)	0.0951		
	miR-224(High vs Low)	3.32(1.15,9.59)	0.0137	2.77(0.95, 8.105)	0.063
Italian set 2	Age (<median vs >median)	0.74(0.32,1.69)	0.4757		
	Gender (Male vs Female)	2.2(0.9,5.36)	0.0744		
	Tumor location (Left vs Right)	1.68(0.73,3.83)	0.2148		
	Tumor stage(III-IV vs I-II)	4.04(1.7,9.57)	0.0006 *	3.89(1.63,9.26)	0.0022 *
	miR-224(High vs Low)	4.41(1.03,18.84)	0.0207 *	4.14(0.96,17.76)	0.056
Romania	Age (<median vs >median)	1.16(0.42,3.2)	0.7733		
	Gender (Male vs Female)	2.44(0.69,8.66)	0.154		
	Tumor location (Left vs Right)	1.3(0.44,3.9)	0.633		
	Tumor stage(III-IV vs I-II)	9.33(2.1,41.54)	0.0004 *	7.47(1.52,36.65)	0.0132 *
	miR-224(High vs Low)	4.25(0.96,18.89)	0.0331 *	1.76(0.36,8.64)	0.4831
Austria	Age (<median vs >median)	0.88(0.51,1.49)	0.6239		
	Gender (Male vs Female)	0.82(0.48,1.43)	0.4909		
	Tumor location (Left vs Right)	0.62(0.36,1.06)	0.0802		
	Tumor Stage(Stage III-IV vs Stage II)	2.89(1.02,8.15)	0.037 *	3.41(1.18,9.83)	0.023 *
	miR-224(High vs Low	2.14(1.21,3.77)	0.0059 *	2.36(1.32,4.21)	0.0037 *
UK	Age (<median vs >median)	1.15(0.31,4.3)	0.8322		
(Overall survival)	Gender (Male vs Female)	0.8(0.21,2.98)	0.7399		
	Tumor location (Left vs Right)	0.98(0.24,3.92)	0.9781		
	Tumor stage(III-IV vs I-II)	2.7(0.68,10.83)	0.1428		
	miR-224(High vs Low)	4.92(1.31,18.46)	0.0181 *		
UK	Age (<median vs >median)	1.54(0.5,4.72)	0.446		
(Metastasis free survival)	Gender (Male vs Female)	0.75(0.25,2.24)	0.6074		
	Tumor location (Left vs Right)	0.6(0.16,2.18)	0.4326		
	Tumor stage(III-IV vs I-II)	2.92(0.93,9.21)	0.0558		
	miR-224(High vs Low)	6.51(1.97,21.51)	0.0008 *		

Table 3: Association of clinical parameters or gene expression with colorectal cancer patient survival

* Statistically significant with p less than 0.05

Figure 1

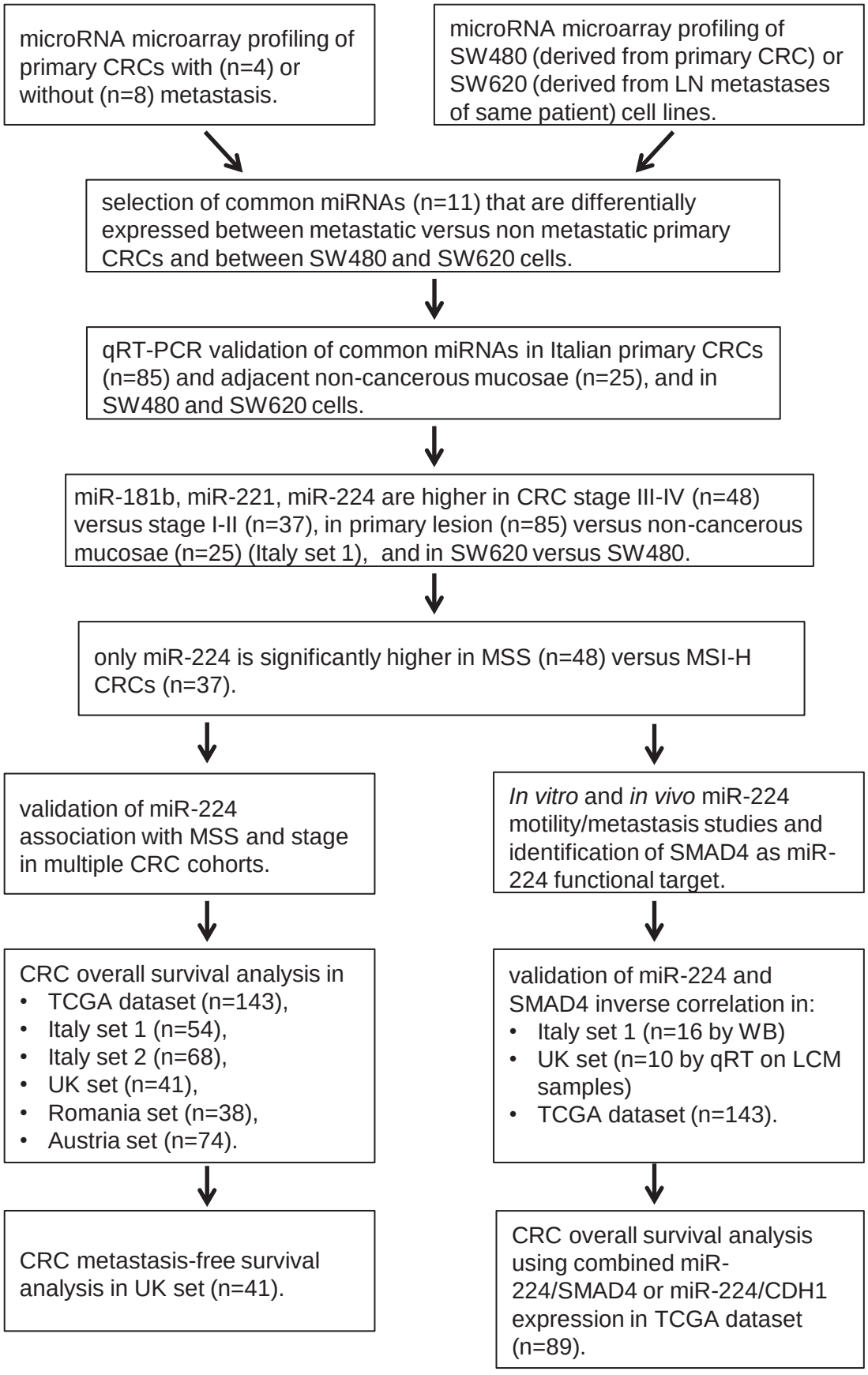

microRNA microarray profiling of primary CRCs with (n=4) or without (n=8) metastasis.
microRNA microarray profiling of SW480 (derived from primary CRC) or SW620 (derived from LN metastases of same patient) cell lines.
selection of common miRNAs (n=11) that are differentially expressed between metastatic versus non metastatic primary CRCs and between SW480 and SW620 cells.
qRT-PCR validation of common miRNAs in Italian primary CRCs (n=85) and adjacent non-cancerous mucosae (n=25), and in SW480 and SW620 cells.
miR-181b, miR-221, miR-224 are higher in CRC stage III-IV (n=48) versus stage I-II (n=37), in primary lesion (n=85) versus non-cancerous mucosae (n=25) (Italy set 1), and in SW620 versus SW480.
only miR-224 is significantly higher in MSS (n=48) versus MSI-H CRCs (n=37).
validation of miR-224 association with MSS and stage in multiple CRC cohorts.
CRC overall survival analysis in
• TCGA dataset (n=143),
• Italy set 1 (n=54),
• Italy set 2 (n=68),
• UK set (n=41),
• Romania set (n=38),
• Austria set (n=74).
CRC metastasis-free survival analysis in UK set (n=41).
In vitro and in vivo miR-224 motility/metastasis studies and identification of SMAD4 as miR-224 functional target.
validation of miR-224 and SMAD4 inverse correlation in:
• Italy set 1 (n=16 by WB)
• UK set (n=10 by qRT on LCM samples)
• TCGA dataset (n=143).
CRC overall survival analysis using combined miR-224/SMAD4 or miR-224/CDH1 expression in TCGA dataset (n=89).

Figure 2

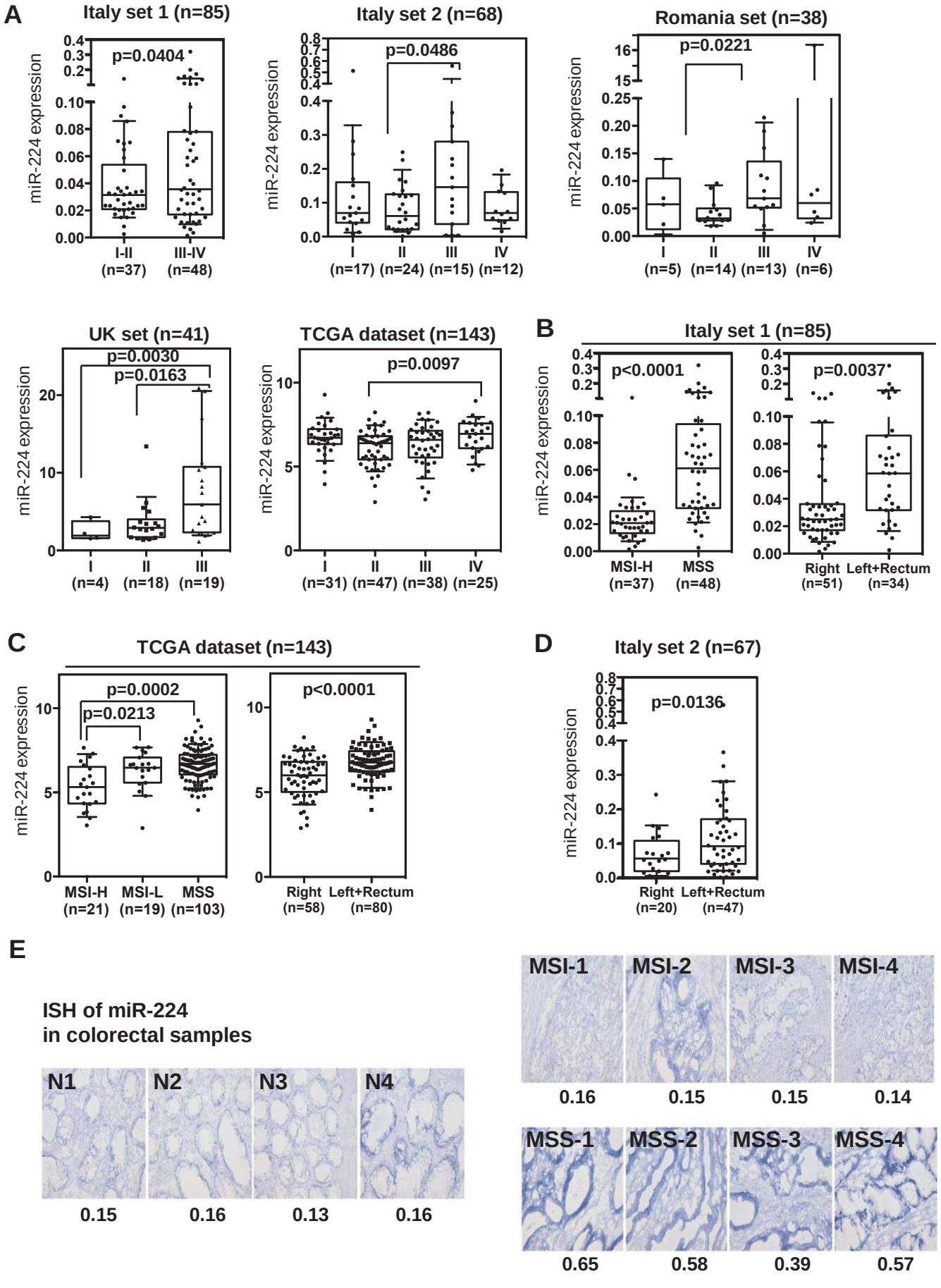

A
Italy set 1 (n=85)
p=0.0404
miR-224 expression
I-II (n=37)
III-IV (n=48)
Italy set 2 (n=68)
p=0.0486
I (n=17)
II (n=24)
III (n=15)
IV (n=12)
Romania set (n=38)
p=0.0221
I (n=5)
II (n=14)
III (n=13)
IV (n=6)
UK set (n=41)
p=0.0030
p=0.0163
I (n=4)
II (n=18)
III (n=19)
TCGA dataset (n=143)
p=0.0097
I (n=31)
II (n=47)
III (n=38)
IV (n=25)
B
Italy set 1 (n=85)
p<0.0001
MSI-H (n=37)
MSS (n=48)
p=0.0037
Right (n=51)
Left+Rectum (n=34)
C
TCGA dataset (n=143)
p=0.0002
p=0.0213
MSI-H (n=21)
MSI-L (n=19)
MSS (n=103)
p<0.0001
Right (n=58)
Left+Rectum (n=80)
D
Italy set 2 (n=67)
p=0.0136
Right (n=20)
Left+Rectum (n=47)
E
ISH of miR-224
in colorectal samples
N1
N2
N3
N4
0.15
0.16
0.13
0.16
MSI-1
MSI-2
MSI-3
MSI-4
0.16
0.15
0.15
0.14
MSS-1
MSS-2
MSS-3
MSS-4
0.65
0.58
0.39
0.57

Figure 3

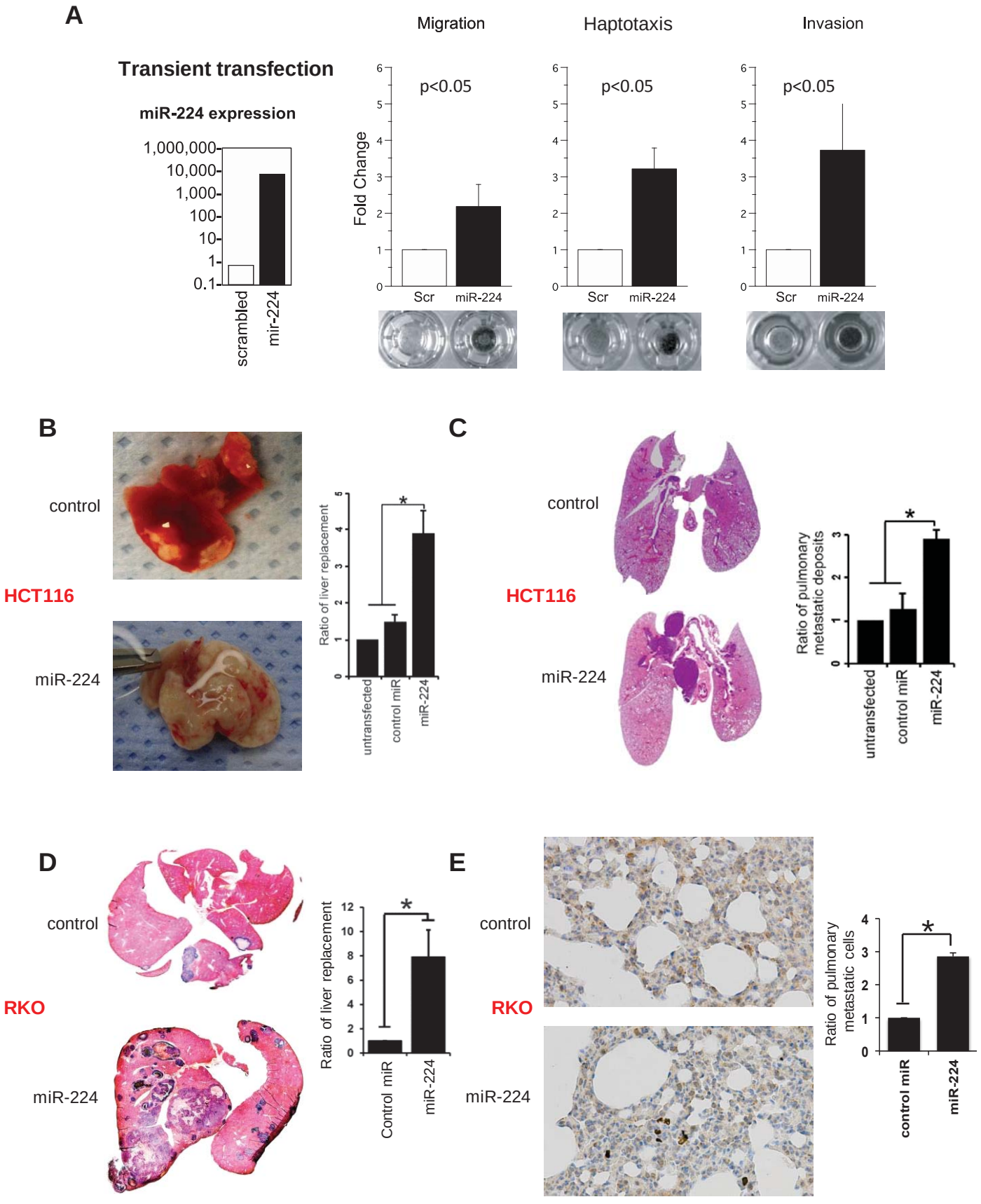

A
Transient transfection
miR-224 expression
1,000,000
10,000
1,000
100
10
1
0.1
scrambled
mir-224
Migration
Haptotaxis
Invasion
p<0.05
p<0.05
p<0.05
Fold Change
Scr
miR-224
Scr
miR-224
Scr
miR-224
B
control
HCT116
miR-224
Ratio of liver replacement
untransfected
control miR
miR-224
C
control
HCT116
miR-224
Ratio of pulmonary metastatic deposits
untransfected
control miR
miR-224
D
control
RKO
miR-224
Ratio of liver replacement
Control miR
miR-224
E
control
RKO
miR-224
Ratio of pulmonary metastatic cells
control miR
miR-224

Figure 4

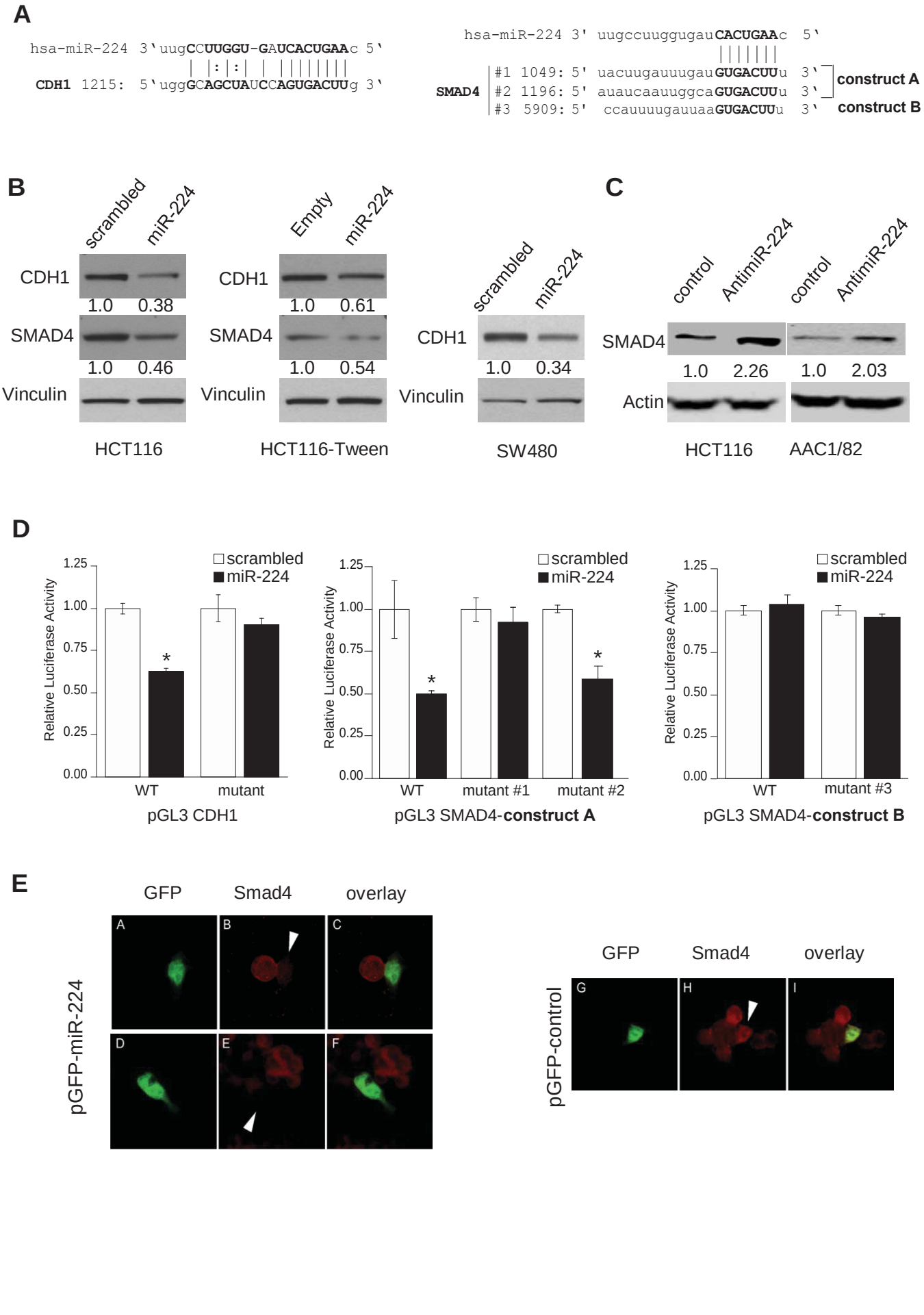
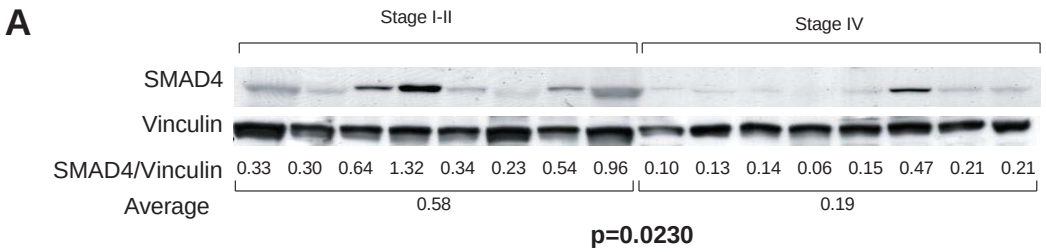
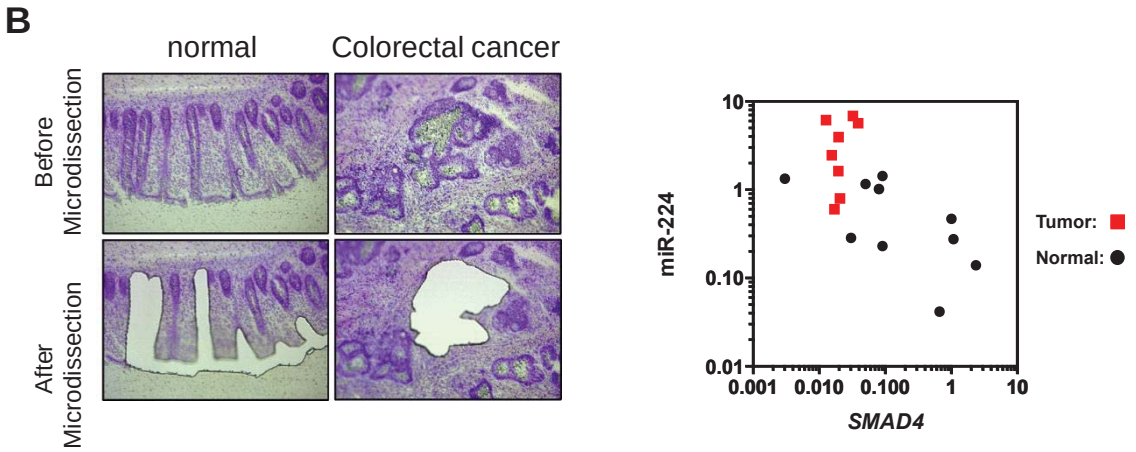
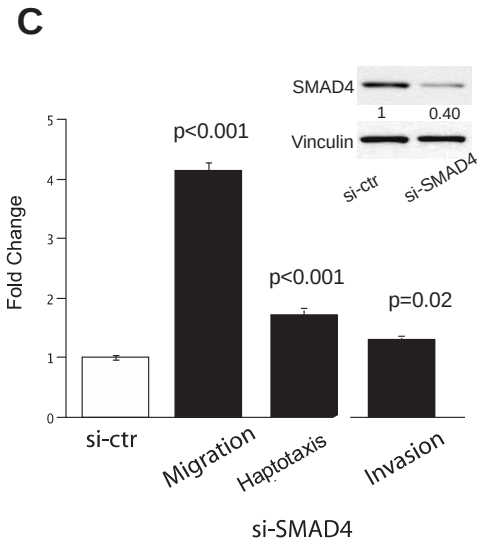

A
hsa-miR-224 3' uugCCUUGGU-GAUCACUGAAc 5'
CDH1 1215: 5' uggGCAGCUAUCCAGUGACUUg 3'
hsa-miR-224 3' uugccuuggugauCACUGAAc 5'
SMAD4
#1 1049: 5' uacuugauuugauGUGACUUu 3'
#2 1196: 5' auaucaauuggcaGUGACUUu 3'
#3 5909: 5' ccauuuugauuaaGUGACUUu 3'
construct A
construct B
B
scrambled
miR-224
Empty
CDH1
SMAD4
Vinculin
1.0 0.38
1.0 0.46
1.0 0.61
1.0 0.54
1.0 0.34
HCT116
HCT116-Tween
SW480
C
control
AntimiR-224
SMAD4
Actin
1.0 2.26
1.0 2.03
HCT116
AAC1/82
D
scrambled
miR-224
Relative Luciferase Activity
1.25
1.00
0.75
0.50
0.25
0.00
WT
mutant
pGL3 CDH1
mutant #1
mutant #2
pGL3 SMAD4-construct A
mutant #3
pGL3 SMAD4-construct B
E
GFP
Smad4
overlay
pGFP-miR-224
pGFP-control
A
B
C
D
E
F
G
H
I

Figure 5


A
Stage I-II
Stage IV
SMAD4
Vinculin
SMAD4/Vinculin 0.33 0.30 0.64 1.32 0.34 0.23 0.54 0.96 0.10 0.13 0.14 0.06 0.15 0.47 0.21 0.21
Average 0.58 0.19
p=0.0230


B
normal
Colorectal cancer
Before Microdissection
After Microdissection
miR-224
SMAD4
Tumor:
Normal:


C
Fold Change
p<0.001
p<0.001
p=0.02
SMAD4
Vinculin
si-ctr
si-SMAD4
si-ctr
Migration
Haptotaxis
Invasion
si-SMAD4

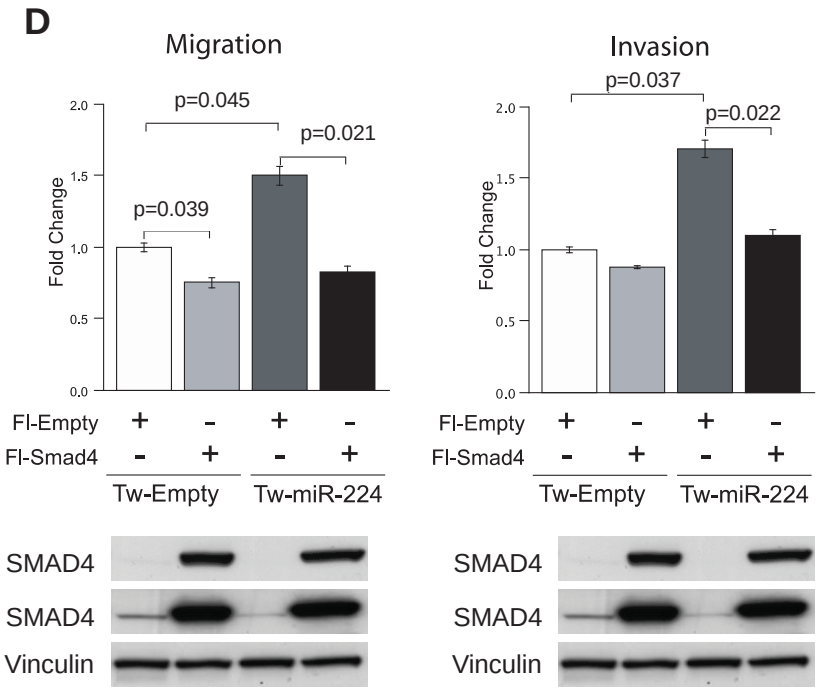
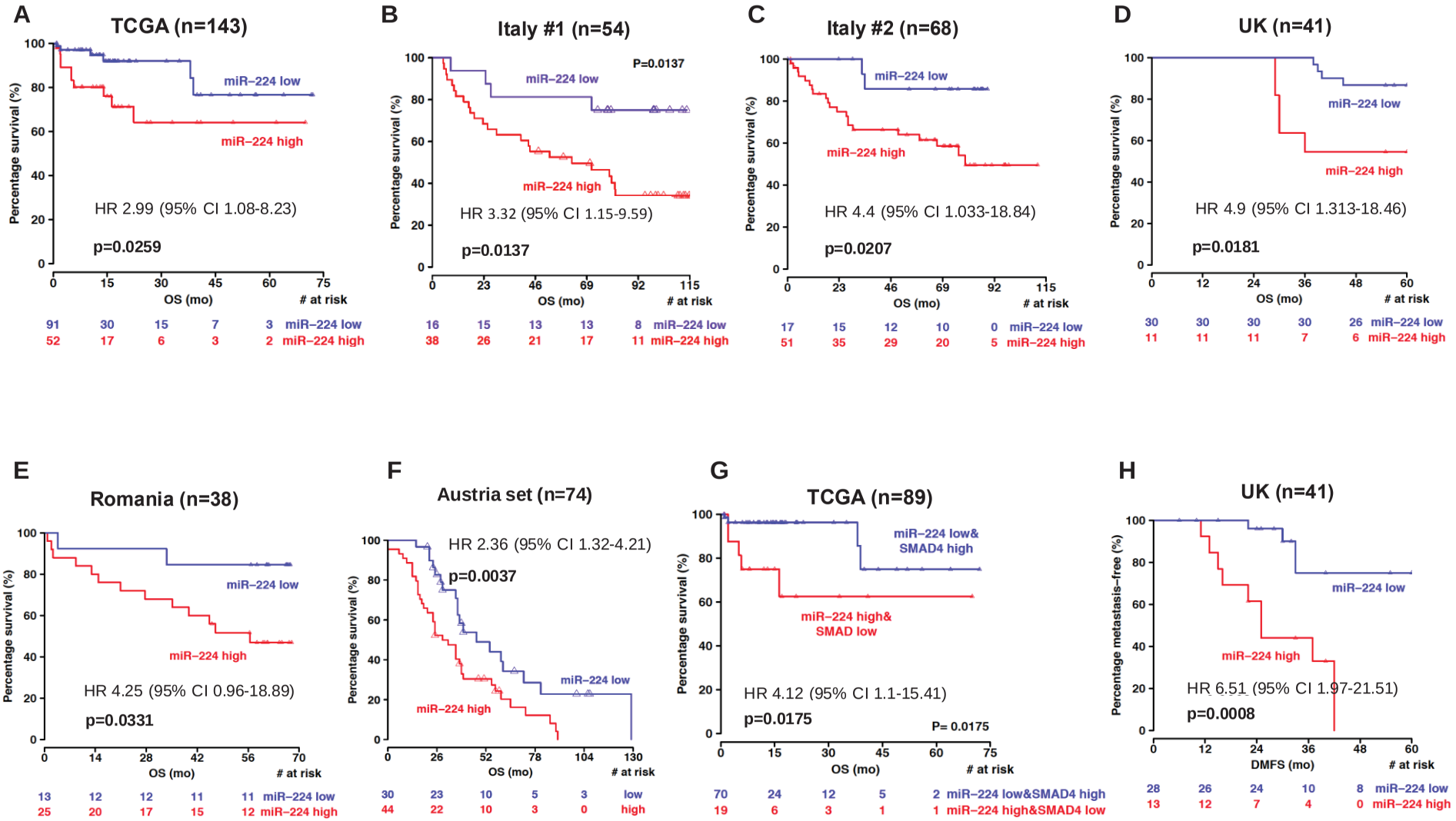

D
Migration
Invasion
p=0.045
p=0.021
p=0.039
p=0.037
p=0.022
Fold Change
Fl-Empty
Fl-Smad4
Tw-Empty
Tw-miR-224
SMAD4
SMAD4
Vinculin

Figure 6


A
TCGA (n=143)
Percentage survival (%)
miR–224 low
miR–224 high
HR 2.99 (95% CI 1.08-8.23)
p=0.0259
OS (mo)
at risk
91 30 15 7 3 miR–224 low
52 17 6 3 2 miR–224 high
B
Italy #1 (n=54)
P=0.0137
HR 3.32 (95% CI 1.15-9.59)
p=0.0137
16 15 13 13 8 miR–224 low
38 26 21 17 11 miR–224 high
C
Italy #2 (n=68)
HR 4.4 (95% CI 1.033-18.84)
p=0.0207
17 15 12 10 0 miR–224 low
51 35 29 20 5 miR–224 high
D
UK (n=41)
HR 4.9 (95% CI 1.313-18.46)
p=0.0181
30 30 30 30 26 miR–224 low
11 11 11 7 6 miR–224 high
E
Romania (n=38)
HR 4.25 (95% CI 0.96-18.89)
p=0.0331
13 12 12 11 11 miR–224 low
25 20 17 15 12 miR–224 high
F
Austria set (n=74)
HR 2.36 (95% CI 1.32-4.21)
p=0.0037
30 23 10 5 3 low
44 22 10 3 0 high
G
TCGA (n=89)
miR–224 low& SMAD4 high
miR–224 high& SMAD low
HR 4.12 (95% CI 1.1-15.41)
p=0.0175
P= 0.0175
70 24 12 5 2 miR–224 low&SMAD4 high
19 6 3 1 1 miR–224 high&SMAD4 low
H
UK (n=41)
Percentage metastasis–free (%)
DMFS (mo)
28 26 24 10 8 miR–224 low
13 12 7 4 0 miR–224 high