

RESEARCH

Open Access



Investigating the relationship of plasma microRNAs and colorectal cancer risk using genetic evidence

Emmanouil Bouras^{1,2*†}, Christos K. Papagiannopoulos^{2†}, Rima Mustafa^{1,3}, Daniel Sobieski⁴, Stephanie L. Schmit^{4,5}, Anna H. Wu⁶, Hermann Brenner^{7,8}, Christopher I. Li⁹, Andrew T. Chan^{10,11,12,13,14,15}, Andrew J. Pellatt¹⁶, Wei Zheng¹⁷, Temitope O. Keku¹⁸, Victor Moreno¹⁹, Caroline Y. Um²⁰, Bethany Van Guelpen^{21,22}, Amanda I. Phipps^{9,23}, Rish K. Pai²⁴, Sarah J. Lewis²⁵, Richard M. Martin^{25,26}, Marc J. Gunter^{1,27}, Ulrike Peters^{9,23}, Abbas Dehghan¹ and Konstantinos K. Tsilidis^{1,2}

Abstract

Background MicroRNAs (miRNAs) are short, single-stranded RNAs that function as post-transcriptional regulators of gene expression. Although circulating miRNAs have been linked to carcinogenesis, they have not yet been systematically investigated in relation to risk of colorectal cancer (CRC).

Methods We used Mendelian randomization (MR) and colocalization analyses to investigate the association of genetically predicted plasma miRNA concentrations (2083 miRNAs in 710 individuals) with risk of CRC (58,221 cases and 67,694 controls). For miRNAs associated with CRC risk, we also investigated their association with circulating plasma proteins (4907 proteins in 35,559 participants), bidirectionally, using MR. We performed pathway enrichment analysis (PEA) to explore downstream molecular pathways.

Results Associations of five miRNAs with CRC were found in MR and supported in colocalization analyses. Specifically, miR-146a-5p, miR-21-5p, and miR-4707-3p were positively, and miR-1908-5p and miR-6810-3p were inversely associated with CRC risk. Several protein associations were found for these miRNAs (range of proteins with $P < 0.05$: 78–796; 211 with FDR < 5%), and 11 pathways were identified in PEA, including regulation of Erb-B2 receptor tyrosine kinase 4 (miR-6810-3p) and insulin-like growth factor pathways (miR-1908-5p).

Conclusions Our results support a potential implication of miR-146a-5p, miR-21-5p, miR-4707-3p, miR-1908-5p, and miR-6810-3p to CRC risk. However, their downstream effects should be elucidated before they can be utilized as preventive targets.

Keywords MicroRNA, MiRNA, Colorectal cancer, Mendelian randomization, Mechanisms, Proteins

[†]Emmanouil Bouras and Christos K. Papagiannopoulos contributed equally to this work.

*Correspondence:

Emmanouil Bouras
ebouras@ic.ac.uk

Full list of author information is available at the end of the article



Background

Colorectal cancer (CRC) is the third most common malignancy worldwide, accounting for over 10% of all cancer cases, and ranking second in mortality in 2020 [1]. Less than 10% of affected individuals carry inherited high-penetrant mutations, with most cases of CRC being sporadic, and tumors demonstrating molecular heterogeneity [2–4]. High-throughput analyses have revealed several biomarkers, including genetic polymorphisms, protein markers, and metabolites, related to CRC development; however, there has been limited high-throughput investigation for other biomarkers, such as microRNAs [5–7].

MicroRNAs (miRNAs) are short, single-stranded, non-coding RNAs that function as post-transcriptional regulators of gene expression [8]. There are over 2000 miRNA-encoding genes in the human genome, many of which are related to biological activities, such as cell growth, differentiation, apoptosis, and senescence, relevant to carcinogenesis [9]. It has also been recently shown that circulating miRNAs contribute to intercellular communication [10]. In vitro and in vivo experimental studies have suggested that several miRNAs play a role in CRC development [11]. Oncogenic miRNAs (onco-miRs), previously reported to be upregulated in CRC, include miR-21, miR-146, and miRNAs of the miR-17 family [11]. On the other hand, tumor-suppressive miRNAs (ts-miRs), found to be depleted in CRC, include let-7, miR-26, and miR-30. Others, such as miR-29, have been suggested to play a dual role in CRC carcinogenesis [11]. MicroRNAs have been shown to regulate the expression of as many as 60% of human protein-coding genes, many of which are key modulators of molecular pathways relevant to CRC carcinogenesis [12, 13].

While the localization of these functions within colorectal tissue is critical to colorectal tumorigenesis, microRNAs identified in the circulation may reflect underlying dysregulation of biological pathways operating at the tissue-specific level. The presence of microRNAs in the circulation may result from passive release due to apoptosis or necrosis, or from active secretion [14]. These circulating microRNAs may originate from colorectal tissue itself or from systemic responses, and their detection in plasma offers a non-invasive means to reflect molecular alterations that drive tumor development within the colon [15].

We are not aware of any published systematic investigation of miRNAs with CRC risk and have conducted here, an in-depth investigation of the association of miRNAs with CRC risk to help identify biological mechanisms related to CRC carcinogenesis and provide potential novel chemo-preventive targets. Additionally, this is the first agnostic investigation of the association of miRNAs

with the human proteome, adding to the in vitro and in vivo evidence for potential molecular links, which could shed light on their exact role in CRC tumorigenesis [12].

In the present study, we used Mendelian randomization (MR) to investigate the association of genetically predicted plasma miRNA concentrations on risk of CRC. Additionally, for those miRNAs that were potentially causally linked to CRC, we explored their effect on the circulating proteome and pertinent molecular pathways.

Methods

We used MR analyses to investigate the association between genetically proxied miRNAs and CRC risk. For miRNAs that were found to alter CRC risk, we performed high-throughput MR analyses to assess their effects on circulating protein concentrations, bidirectionally, and pathway enrichment analysis, to further delineate potential mechanisms of action. Stratified analyses were also performed by sex and anatomical location of tumor to investigate homogeneity in the associations. An overview of the study design is shown in Fig. 1.

Data sources

GWAS of circulating plasma miRNAs

A genome-wide association study (GWAS) of 710 healthy, unrelated, weight-stable, European ancestry individuals with obesity was used to extract summary genetic association estimates for 2083 circulating plasma miRNAs that were used in our main analyses. Details on study participants, data collection, and quantification were presented previously [16]. In secondary/replication analyses, we used genetic association estimates for 5269 microRNA expression quantitative trait loci (miR-eQTLs) from a Framingham Heart Study (FHS) GWAS ($n=5239$ individuals) study. Detailed information about processing of genotype data and study design can be found elsewhere [17].

GWAS of colorectal cancer

Summary genetic association estimates for CRC and subtypes (colon and rectal cancer, CRC in male and female, and early-onset CRC), in up to 58,221 cases and 67,694 controls, were obtained from a GWAS meta-analysis of the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), Colorectal Transdisciplinary Study (CORECT), and Colon Cancer Family Registry (CCFR) genetic consortia [18].

GWAS of circulating plasma proteins

A GWAS of 35,559 Icelanders was used to extract summary genetic association estimates for circulating plasma proteins [17]. The GWAS provides estimates for 4907

GWAS sources

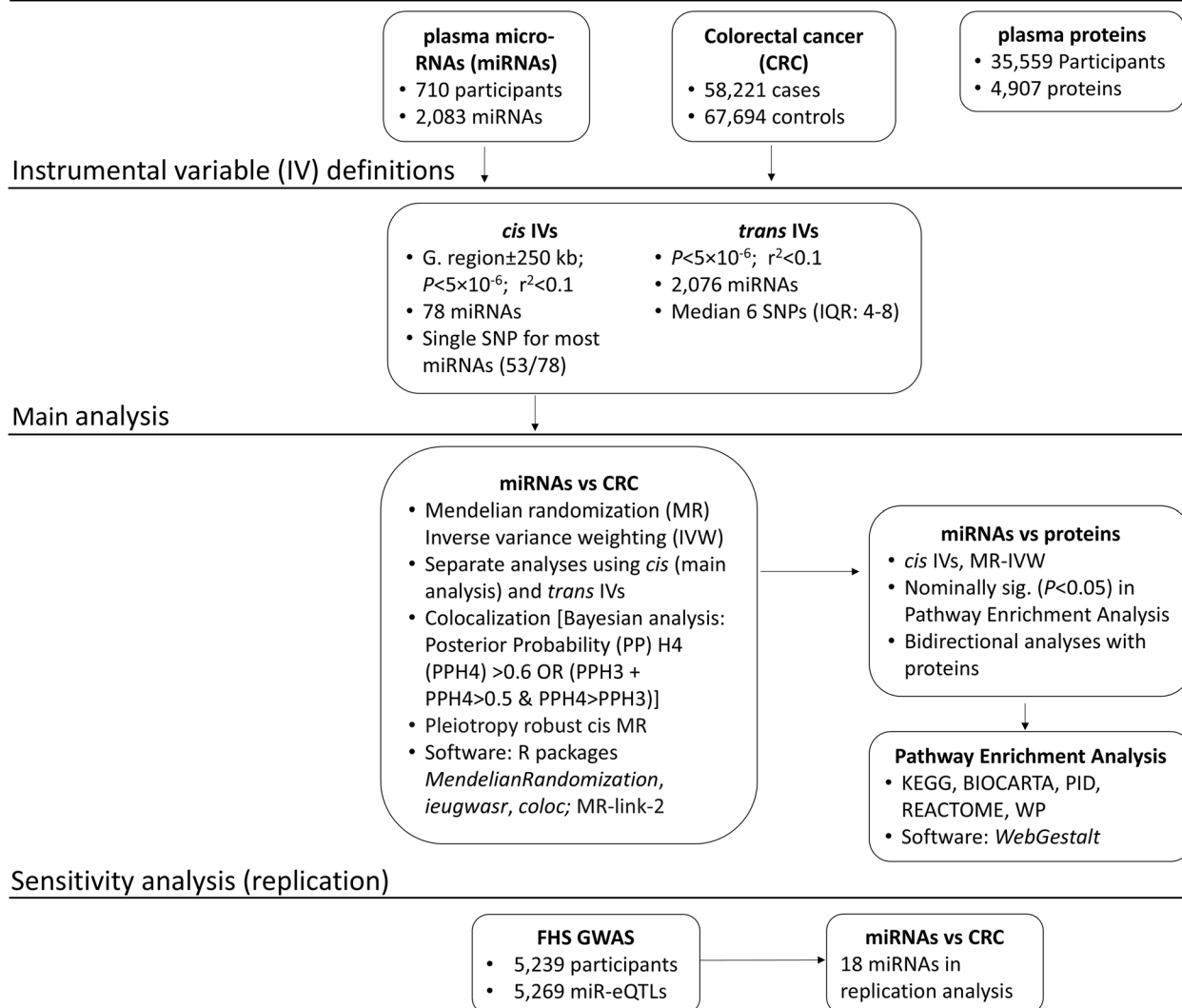


Fig. 1 Overview of the study design. Abbreviations: CRC, colorectal cancer; FDR, false discovery rate; FHS, Framingham Heart Study; GWAS, genome-wide association study; IV, instrumental variable; miRNA, microRNA; eQTL, expression quantitative trait loci; SNP, single nucleotide polymorphism

proteins quantified using an aptamer-based SomaScan assay. Details on processing of genotype data and study design is described elsewhere [19].

Statistical analysis

Genetic instrument definitions

To minimize the possibility of horizontal pleiotropy (i.e., when genetic variants affect CRC risk via pathways unrelated to the miRNA under investigation), we used *cis* instrument definitions (i.e., proximal to pertinent genetic regions rather than from throughout the genome) as our main analysis. Specifically, to develop genetic instrumental variables (IVs) for miRNAs, we first selected SNPs

associated with circulating plasma miRNA concentrations ($P < 5 \times 10^{-6}$), located within ± 250 kb of pertinent genetic loci of miRNAs, that are weakly correlated ($r^2 < 0.1$). Genomic coordinates of miRNA regions were extracted from MiRbase v22 [9]. Second, we developed IVs, using the above criteria, but without restricting to pertinent genetic regions (*trans*-gene IVs), for comparison, and to further identify potential causal associations that were not captured using the *cis* IV definitions. In sensitivity analyses, we applied a more stringent threshold of $P < 5 \times 10^{-8}$ for defining genetic IVs. SNPs with a minor allele frequency (MAF) < 0.01 and palindromic SNPs were removed [20].

For the bidirectional protein to miRNA analyses, we used *cis* IV definitions selecting independent genetic variants ($r^2 < 0.001$) within 1 Mb of each protein-coding gene associated with plasma protein concentrations (P value $< 5 \times 10^{-8}$).

Mendelian randomization

The ratio estimate was used to obtain causal estimates for miRNAs with a single SNP, and the inverse variance weighting (IVW) estimate for miRNAs with ≥ 2 SNPs, accounting for the weak LD among variants [21].

MR can generate unbiased estimates of causal effects of exposures on disease outcomes provided its assumptions are met [22]. Under the three core MR assumptions, for the selected genetic variants used as IVs to be valid instruments they should: (i) be strongly associated with the circulating miRNAs (relevance), (ii) be independent of any potential confounding variable of the circulating miRNAs-cancer association (independence), and (iii) affect CRC risk only through the circulating miRNAs being instrumented (exclusion restriction). To ensure that the first assumption is met, we used variants with a P value $< 5 \times 10^{-6}$ and used only variants with a F -statistic (a measure of instrument strength) > 10 .

The presence of horizontal pleiotropy is the most common reason for violation of the third MR assumption. Using *cis* instruments, we minimized the possibility for horizontal pleiotropy since they influence gene expression in their immediate genomic region and directly affect miRNA rather than multiple, unrelated traits. We employed the MR-link-2 method [23], to estimate the extent of the pleiotropy in the identified loci. In brief, MR-link-2 is a likelihood function that uses the exposure and the outcome summary genetic association estimates in a region, combined with a reference linkage disequilibrium (LD) matrix, and tests for two parameters using a likelihood ratio test: the causal effect estimate ($\hat{\alpha}$) and the remaining horizontal pleiotropic variance (σ_y).

We also utilized the Phenoscanner database to look for previously reported associations of the selected SNPs [24, 25]. Additionally, to test the second and third MR assumptions, we conducted sensitivity analyses using robust MR methods that make different IV assumptions allowing the inclusion of pleiotropic variants, namely the weighted median (WMe) [26], MR-Egger [27, 28], and weighted mode (Wmo) [29], where there were ≥ 3 SNPs available (≥ 2 for Wmo).

We calculated the Benjamini–Hochberg false discovery rate (FDR) to account for the multiple comparisons (of *trans*-defined miRNAs on CRC risk and on proteins) [21].

The effects of genetically proxied miRNAs with CRC risk, which were confirmed in analyses using *cis* IVs and

were supported in colocalization analyses, were considered robust and included in subsequent analyses with plasma proteins (using *cis* IVs).

We considered miRNAs to be associated with individual proteins when FDR $< 5\%$. We also applied an alternative definition: nominal miRNA-protein associations (MR-IVW P value < 0.05) with high predicted binding affinity (score > 50) to relevant genes, in order to highlight suggestively functional miRNA targets. The predicted affinity was evaluated based on a publicly available miRNA target prediction model that was originally developed to identify features that are characteristic of target downregulation and target binding [13]. Nominal associations of miRNAs with proteins (MR-IVW P value < 0.05) were further explored in pathway enrichment analyses.

Additionally, we performed analyses of plasma proteins on miRNAs, to explore potential bidirectional associations and highlight proteins linked to the highlighted miRNAs.

In addition, we investigated the potential reverse association between genetic liability to risk of developing CRC and plasma miRNA concentrations (bidirectional MR), for all miRNAs robustly associated with CRC risk, using MR analyses.

Colocalization analyses

We employed a Bayesian framework for colocalization analysis proposed by Giambartolomei et al., to examine confounding by LD in the associations of miRNAs with CRC risk [30]. The algorithm calculates posterior probabilities (PP) of five different hypotheses based on causal variant configurations: H0 (no causal variant), H1 (causal variant for trait 1 only), H2 (causal variant for trait 2 only), H3 (two distinct causal variants), and H4 (one common causal variant). $PPH4 > 60\%$, or $[(PPH3 + PPH4) > 50\% \text{ and } PPH4 > PPH3]$, was considered evidence to support the presence of a shared causal variant between miRNAs and outcomes [31].

Pathway enrichment analysis

To explore the proteome-wide associations of the miRNAs that were associated with CRC risk, we used the list of the protein-coding genes of the associated proteins (MR-IVW $P < 0.05$) and performed pathway enrichment analysis (PEA), using the 4907 protein-coding genes as background.

We conducted PEA using a WEB-based Gene Set Analysis Toolkit (WebGestalt) and explored enriched pathways in the BIOCARTA, Kyoto Encyclopedia of Genes and Genomes (KEGG), Pathway Interaction Database (PID), REACTOME, and WikiPathways (WP) gene-sets [32]. The model parameters that were used are described in Additional file 1: Table S1.

Expression profiles of miRNAs

We evaluated the expression profiles of the miRNAs that were associated with CRC risk using two publicly available platforms. The first was miRNATissueAtlas 2025, from which we used expression data of 46,997 tissue samples across 74 organs, including 1327 bowel samples [33]. We used this database to evaluate miRNA expression in the healthy colon tissue and compare it with plasma expression. The second was miRNASNP-v3, from which we used data on the expression of miRNAs and their target genes in 33 cancer types from The Cancer Genome Atlas (TCGA), and correlation coefficients between expression of miRNAs and their target genes across cancer types [34]. We also used the latter platform to investigate the correlation between miRNA expression and drug (small compound) sensitivity, measured using the half-maximal growth inhibitory concentration (GI50) in the National Cancer Institute (NCI) NCI-60 cancer cell line pharmacogenomic database [35]. The correlations were estimated using Pearson's correlation coefficient between the GI50 values and the miRNA expression levels, and coefficients with an FDR < 0.05 were considered significant. The database contained information on GI50 of 18,724 compounds and expression profiles of 335 miRNAs, at the time of assessment.

Secondary analyses

Because the main analysis genetic association estimates for miRNAs were obtained from a population of individuals with obesity, we used genetic association estimates for 5269 miR-eQTLs from a FHS GWAS study, to explore homogeneity in the genetic associations.

All analyses were performed using R version 4.3.1 (2023-06-16 ucrt) and the “*MendelianRandomization*,” “*iiegwasr*,” “*bigsnp*,” and “*coloc*” packages [36].

Results

Instrument characteristics

We found *trans* instruments for 2076 miRNAs with a median of 6 SNPs (IQR: 4 to 8 SNPs) per miRNA, and *cis* instruments for 78 of the 2076 miRNAs, of which 68% (53/78) were comprised of a single SNP. The median *F*-statistic was 23 (IQR: 22–25) across *trans* IVs and 36 (IQR: 27–57) across *cis* IVs. Details of the miRNAs and the SNPs that were used as IVs are presented in Additional file 1: Table S2. In sensitivity analysis using the *P* value threshold of 5×10^{-8} , a subset of 348 *trans*-defined and 51 *cis*-defined miRNAs could be proxied. To maximize discovery potential, we used the less stringent threshold of 5×10^{-6} in our analysis.

Allele frequencies across the 1737 rsIDs shared between the microRNA and protein GWAS datasets were highly similar ($\rho = 0.997$; Additional file 2: Fig. S1).

Evaluating the association of miRNAs with colorectal cancer

Eight nominal associations with CRC were found using *cis* IVs (Additional file 1: Table S3). One hundred thirty-seven nominal associations with CRC were found using *trans* IVs, only one of which remained after multiple-testing correction (miR-1908-5p) (Fig. 2; Additional file 1: Table S4). Six associations were replicated using both *cis* and *trans* definitions (Fig. 2). When we compared the MR estimates from *cis* and *trans*-defined IVs (regardless of significance), 74% of the associations (58/78) agreed in terms of direction of effects with a Pearson's correlation coefficient of 0.73 (*P* value < 0.001) (Additional file 2: Fig. S2). When we included only the eight miRNAs that showed nominal associations with CRC risk using the *cis* IVs, the correlation was 0.83 (*P* value = 0.01) (Additional file 2: Fig. S3). Colocalization analyses were conducted for all the associations that were found using the *cis* IVs and provided evidence to support the presence of shared causal variants with CRC risk for five miRNAs, namely miR-146a-5p, miR-21-5p, miR-4707-3p, miR-1908-5p, and miR-6810-3p (Additional file 1: Table S5; Additional file 2: Figs. S4–S8). These five miRNAs were considered to be robustly associated with CRC risk (Fig. 3). Four of these associations (miR-146a-5p, miR-21-5p, miR-4707-3p, and miR-1908-5p) were replicated in sensitivity analyses using a more stringent *P* value threshold for defining genetic instruments (*P* value < 5×10^{-8}) (Additional file 1: Table S6). The miRNA miR-6810-3p could not be tested in the MR analysis due to a lack of instruments. The associations were largely homogeneous in the stratified analyses, by sex, in colon and rectal cancer, and early-onset CRC (Additional file 2: Fig. S9).

Pleiotropy-robust MR analyses using the MR-link-2 method further supported the associations identified in the main MR results, showing concordant associations (Additional file 1: Table S7). However, the causal effect parameter P_a for miR-146a-5p and miR-21-5p did not reach statistical significance in MR-link-2 ($P_a = 0.15$ and 0.21, respectively), possibly due to limited statistical power. Notably, an association for miR-146a-5p was identified in a *trans* locus (chr1:54,791,012–55791012), where the strongest variant is intronic to ciliary microtubule associated protein 2 (*CIMAP2*). Only limited evidence of pleiotropy was observed in the *cis*-regions included in the main analyses, with a significant pleiotropy parameter ($P_\sigma < 0.05$) detected only for the *cis*-region of miR-4707-3p—suggesting a potential pleiotropic effect in addition to the potentially causal association (Additional file 1: Table S7). In assessing pleiotropy for the SNPs that were used as *cis* IVs, using *Phenoscan*, we found no major pleiotropic pathways, except for rs174561 (that was used as single IV for miR-1908-5p), which was associated

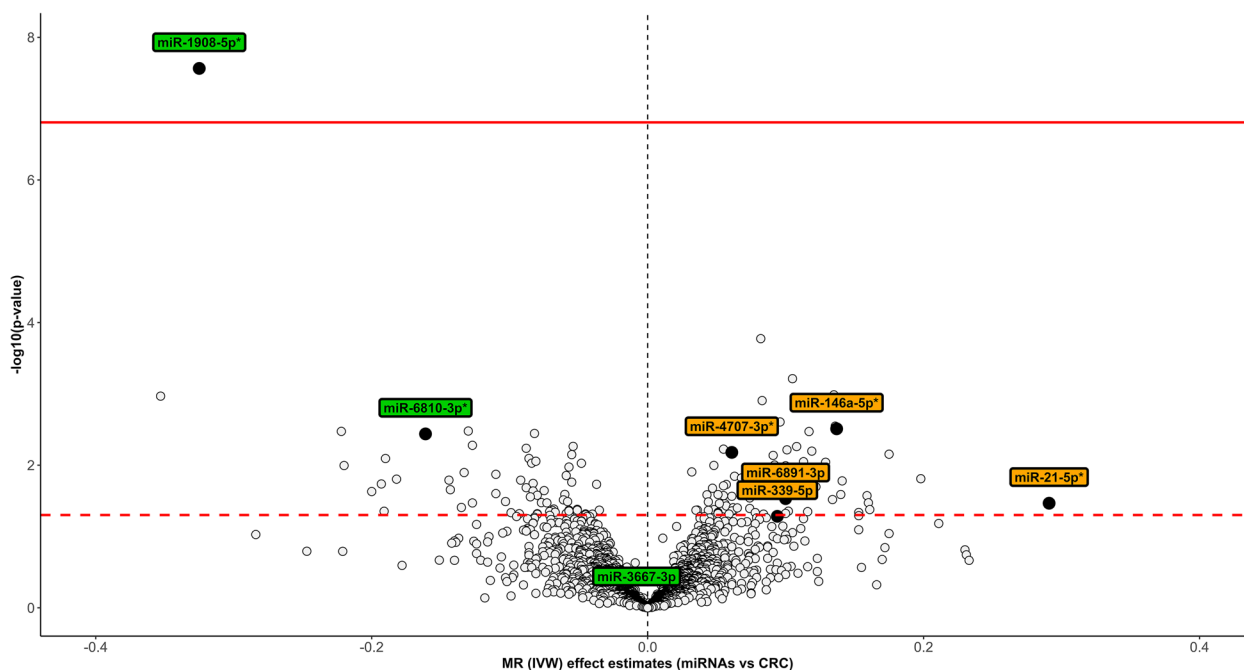


Fig. 2 Volcano plot of the associations between microRNAs (miRNAs) and colorectal cancer (CRC) risk. The X-axis shows the Mendelian randomization (MR) inverse variance weighting (IVW) estimates for the associations between miRNAs and CRC risk, using trans instruments, and the Y-axis the pertinent $-\log_{10} P$ values. The threshold of significance is indicated by the two lines (dashed line: P value < 0.05; solid line: FDR < 5%). Black colored labels represent miRNAs that were associated with CRC risk using cis instruments, whereas the asterisk (shown in the labels) indicates that the association was supported in colocalization analysis. Abbreviations: CRC, colorectal cancer; FDR, false discovery rate; IVW, inverse variance weighting; MR, Mendelian randomization; miRNAs, microRNAs

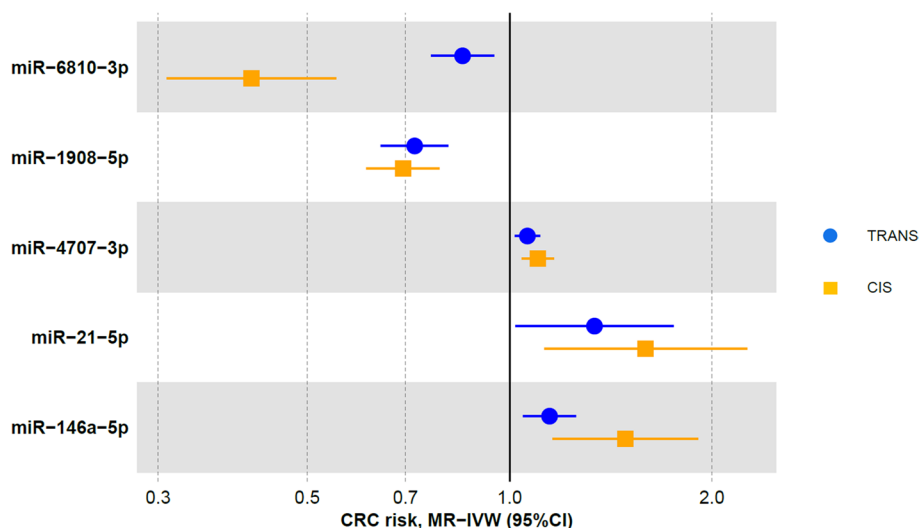


Fig. 3 Forest plot of the associations of microRNAs (miRNAs) with colorectal cancer (CRC) risk. The associations of five miRNAs that showed significant associations with CRC risk in the Mendelian randomization (MR)-inverse variance weighting (IVW) analyses and supported in colocalization analyses are presented. Different shapes correspond to different instrumental variable definitions. Abbreviations: CRC, colorectal cancer; IVW, inverse variance weighting; MR, Mendelian randomization; miRNAs, microRNAs

with plasma lipids and inflammatory bowel disease, and rs1473901 (single IV for miR-6810-3p) which was associated with body composition related phenotypes (Additional file 1: Table S8).

Instruments for 18 cis-defined miRNAs were available from both the main GWAS and the FHS GWAS, the associations of which with CRC risk were qualitatively consistent and the MR-IVW of the two GWASs

(regardless of significance) were moderately correlated ($r^2=0.43$), providing some evidence of homogeneity in the genetic associations (Fig. 4). However, none of the five miRNAs that were highlighted in our analysis was available in the FHS GWAS.

We found little evidence for bidirectional associations between genetic liability to risk of developing CRC and plasma miRNA concentrations for any of the five miRNAs (Additional file 1: Table S9).

Evaluating the association of miRNAs with proteins

The potential role of the five miRNAs that were robustly associated with in CRC risk was explored using high-throughput MR analyses, looking at their effect on 4907 plasma proteins. Most nominal associations were found for miR-1908-5p (796/4907), followed by miR-4707-3p (519/4907), miR-21-5p (249/4907), miR-146a-5p (246/4907), and miR-6810-3p (78/4907) (Fig. 5; Additional file 1: Table S10). Among those proteins, 86 were associated (P value < 0.05) with CRC risk (Additional file 1: Table S11). After correcting for multiple comparisons, 203 associations remained for miR-1908-5p, six for miR-6810-3p, two for miR-21-5p, and none for miR-4707-3p and miR-146a-5p (Fig. 5; Additional file 1: Table S10). Ten suggestively functional miRNA protein targets (genes/proteins having a high predicted affinity score with pertinent miRNAs and nominally associated in MR) were found for miR-21-5p, seven for miR-1908-5p, three for miR-146a-5p,

and one each for miR-4707-3p and miR-6810-3p (Fig. 5; Additional file 1: Table S12). After correcting for multiple testing ($FDR < 0.05$), we found no evidence of bidirectional associations between proteins and plasma miRNAs (Additional file 1: Table S13). Only one protein—nuclear factor kappa B subunit 1 (*NFKB1*)—was found to be (unidirectionally) associated with plasma miRNAs at $FDR < 0.05$, specifically with miR-146a-5p.

Pathway analysis

Pathway analysis showed four REACTOME pathways enriched for miR-146a-5p [association of TriC/CCT with target proteins during biosynthesis; chaperonin-mediated protein folding; protein folding; cooperation of PDCL (PhLP1) and TRiC/CCT in G-protein beta folding], three pathways for miR-6810-3p [downregulation of ERBB4 signaling; neuexins and neuroligins; striated muscle contraction], one for miR-1908-5p [regulation of insulin-like growth factor (IGF) transport and uptake by insulin-like growth factor binding proteins (IGFBPs)], and one Wikipathway each for miR-4707-3p [amino acid metabolism], miR-1908-5p [striated muscle contraction pathway], and miR-6810-3p [striated muscle contraction pathway] (Additional file 1: Table S14).

Expression profiles of miRNAs

All five miRNAs with robust associations with CRC risk showed expression in the healthy colon tissue (Additional

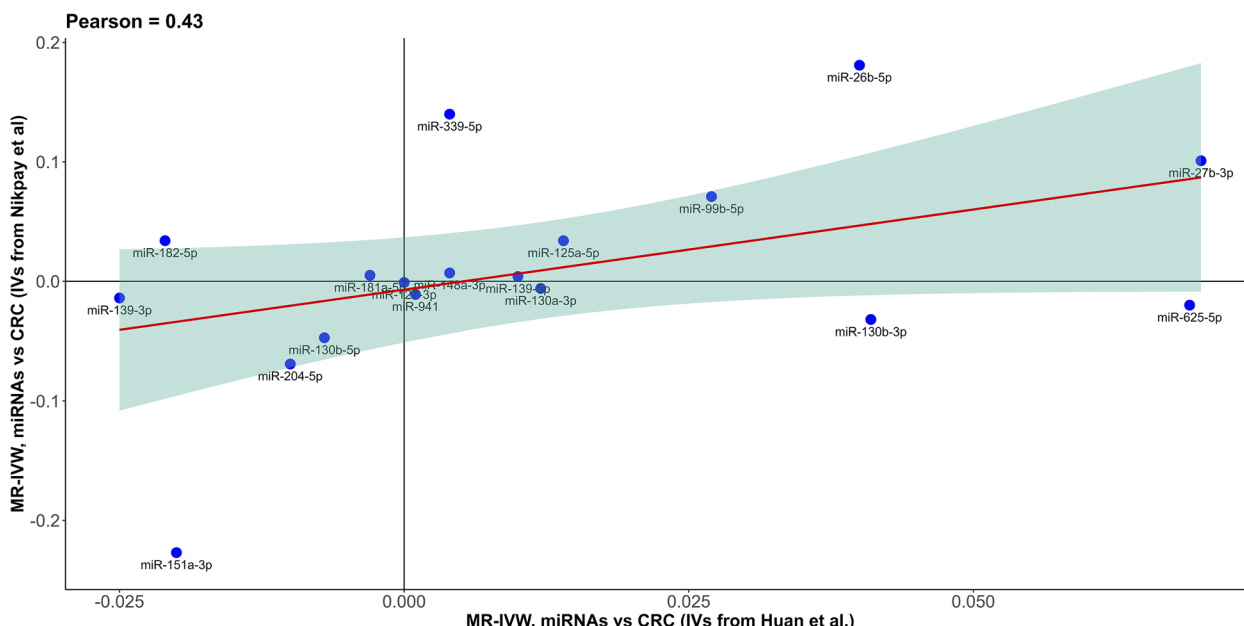


Fig. 4 Correlation of the Mendelian randomization (MR) estimates using genetic summary data from different genome-wide association studies. MR-inverse variance weighting (IVW) estimates from the main analysis genome-wide association study (GWAS, Nikpay et al.) versus MR-IVW estimates from the sensitivity analysis GWAS (Huan et al.), using cis instruments, for 18 microRNAs (miRNAs) that were available in both resources. Abbreviations: CRC, colorectal cancer; GWAS, genome-wide association study; IVW, inverse variance weighting; IVs, instrumental variables; MR, Mendelian randomization; miRNAs, microRNAs

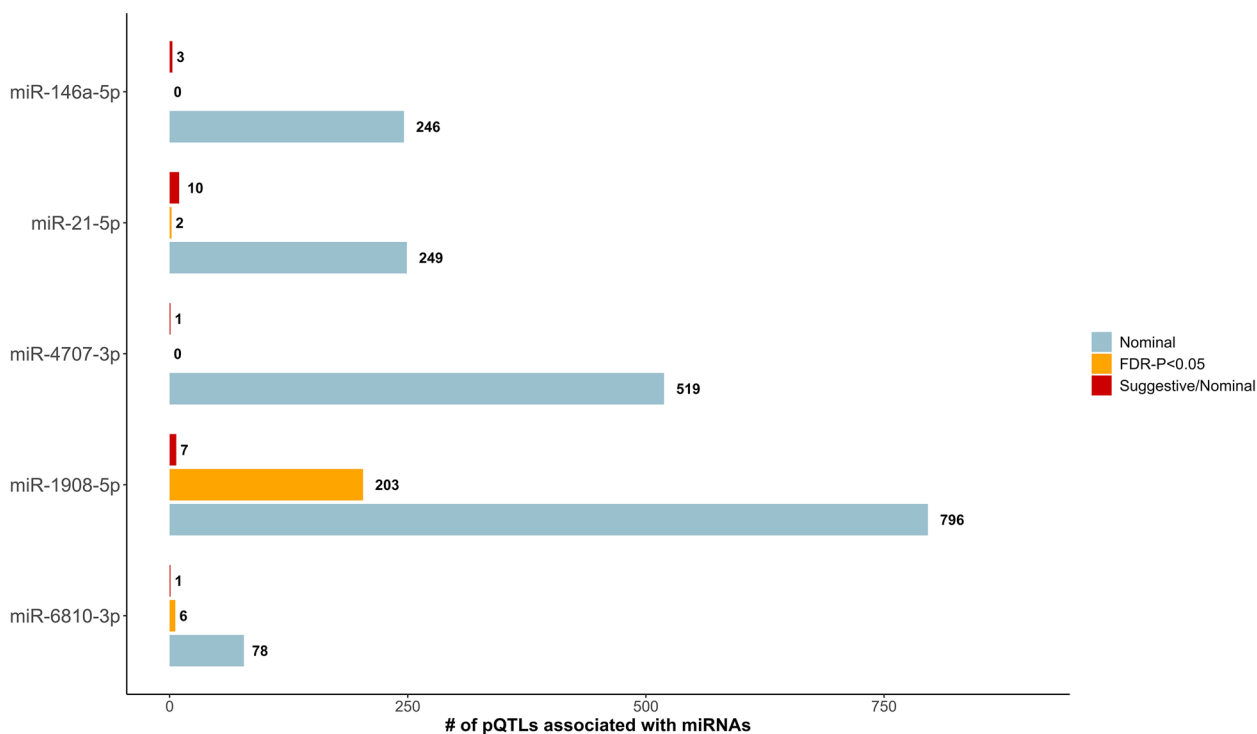


Fig. 5 Number of proteins associated with each microRNA (miRNA). Associations include nominal miRNA to protein [Mendelian randomization (MR)-inverse variance weighting (IVW) P value < 0.05], associations significant based on False discovery rate (FDR) < 5%, and suggestive/nominal associations of miRNAs with proteins (MR-IVW P value < 0.05), which had a high predicted affinity (score > 50) with pertinent genes, based on a publicly available miRNA target prediction model. Abbreviations: FDR, false discovery rate; IVW, inverse variance weighting; MR, Mendelian randomization

file 1: Table S15a; Additional file 2: Fig. S10). Across all identified miRNAs, there was a good correlation of colon versus plasma tissue expression ($\rho = 0.86$; P value = 0.006), across healthy tissue samples (Additional file 2: Fig. S11).

All five miRNAs were expressed in colon adenocarcinoma tissues (Additional file 1: Table S15b). Expression of miR-21-5p and miR-146a-5p correlated (Pearson’s correlation P value < 0.05) with the expression (mRNA) of 799 and 516 unique genes, respectively, across cancer tissues (Additional file 1: Table S16).

Additionally, the expression of miR-21-5p and miR-146a-5p correlated (FDR < 5%) with sensitivity to 706 and 46 drug/small molecule compounds, respectively (Additional file 1: Table S17). No data was available for the rest of the associated miRNAs.

Discussion

We performed high-throughput MR analyses to agnostically investigate the potential links between circulating miRNAs and risk of CRC and further explored miRNA-associated plasma proteins. Our study provides evidence, from the MR and colocalization analyses, to support that genetically predicted plasma miR-146a-5p, miR-21-5p,

and miR-4707-3p (acting as oncogenic miRNAs), and miR-1908-5p and miR-6810-3p (acting as tumor suppressor miRNAs), could serve as potential novel targets for CRC chemoprevention (e.g., by pharmacological agents, or lifestyle modification). However, several downstream protein targets and relevant pathways were also identified for these miRNAs, which should be thoroughly characterized before they can be used in clinical practice.

Genetically predicted miR-146a-5p was positively associated with CRC risk in our analyses, in accordance with previous in vitro studies where its role as onco-miR has been demonstrated [11, 37]. Such studies suggest that miR-146a-5p is involved in the regulation of intestinal stem cells and implicate targets within the *Notch* and *Wnt*, as well as the nuclear factor *NF-κB* signaling pathways as downstream mediators [37, 38]. We found a suggestive association between miR-146a-5p and zinc and ring finger 3 (*ZNRF3*), a protein that acts as a negative regulator of the *Wnt* signaling pathway, which aligns with the above hypothesis. Interestingly, we found a positive association between genetically predicted plasma NFKB1 and miR-146a-5p, supporting the implication of the two biomarkers in these shared biological pathways. Previous

studies have also implicated miR-146a as a negative regulator of toll-like receptor (TLR)-mediated inflammation [39]. A recent in vivo study found that miR-146a acted as a negative regulator of colonic inflammation and associated tumorigenesis by inhibiting IL-17 responses [40]. Major targets were *RIPK2* (modulating the downstream *NOD2* signaling pathway) and *TRAF6* (modulating the MAPK and NF κ B pathways). Additionally, it has been suggested that miR-146a modulates prostaglandin E2 (PGE2) in tumor cells within intestinal epithelial cells [41]. In line with the latter observation, we found a suggestive association between miR-146a-5p and prostaglandin F2 receptor inhibitor (*PDGFRN*), which is involved in the *prostaglandin synthesis and regulation* pathway. Our results also suggest an association between miR-146a-5p and fucosyltransferase 9 (*FUT9*), a protein which is implicated in glycosphingolipid biosynthesis. Previous experimental studies have demonstrated that miR-146a is involved in the conversion of erythropoiesis to myelopoiesis that occurs due to inflammatory signaling mediated by sphingolipids, which leads to the inhibition of autophagy, in human hematopoietic stem/progenitor cells [42].

Our study provides evidence to support the role of miR-21-5p as an onco-miR for CRC risk, in line with previously published studies [11]. Mechanistically, it has been reported that miR-21-5p might be associated with migration, invasion, angiogenesis and metastasis, promoting epithelial mesenchymal transition (EMT), hyperactivation of the PI3K-AKT [43], and the Wnt signaling pathway [44], as well as regulation of inflammatory signaling pathways [cyclooxygenase (COX)-2 inflammation pathway] [45]. We found little evidence to support the above hypotheses at the pathway-level; however, many of the proteins associated with miR-21-5p are involved in immunomodulatory processes [such as C-C motif chemokine ligand 3 (*CCL3*), interleukin 1B (*IL1B*), *IL6R*, and vascular cell adhesion molecule 1 (*VCAMI*)], participate in the organization of the extracellular matrix (such as several collagen-type proteins and matrix metalloproteinases), and apoptosis [death associated protein kinase 1 (*DAPK1*), mitogen-activated protein kinase 8 (*MAPK8*), tumor protein P63 (*TP63*)]. Furthermore, expression of miR-21-5p correlated with sensitivity of over 700 small molecule compounds; however, the potential utility of this biomarker as a pharmacological target for CRC is yet to be explored [46].

We found a potential association between genetically predicted miR-4707-3p and CRC risk. In a murine study, miR-4707-3p was found to interact with DANCR to regulate the expression of FOXC2 oncogene, in a zinc finger protein 750 (*ZNF750*) dependent manner, which affected esophageal squamous cell carcinoma angiogenesis [47].

With regard to CRC, previous studies have shown that DANCR might act as an oncogenic long non-coding RNA affecting tumor progression and FOXC2 [48] has been implicated as an oncogene promoting tumor invasion and metastasis [49]. We found little evidence to support the above mechanisms; however, we found a suggestive association between miR-4707-3p with glutathione-disulfide reductase (GSR), suggesting a potential role in antioxidant defense.

The results of our study suggest a potential association between miR-1908-5p and CRC risk. miR-1908-5p has been previously associated with cancer outcomes, such as non-small cell lung cancer, prostate, breast, and epithelial ovarian cancer in vitro and in vivo studies; however, evidence for CRC was limited [50]. These experimental studies have shown that miR-1908-5p affects proliferation through activating downstream pathways such as the PI3K/AKT/mTOR. In support of the above observations, the high-throughput MR analyses that we performed showed that miR-1908 was associated with several proteins involved in the PI3K/AKT/mTOR pathway, such as Klotho (*KL*), MET proto-oncogene, receptor tyrosine kinase (*MET*), and neuregulin 1 (*NRG1*). Additionally, miR-1908-5p was associated with fatty acid synthase (*FAS*) and several apolipoproteins (e.g., apolipoprotein B, F, and L1), and enrichment analyses showed that miR-1908-5p was associated with proteins related to the regulation of IGF transport and uptake via IGF-BPs. These results suggest a potential role of miR-1908-5p in metabolic regulation that might be associated with CRC. MicroRNA miR-1908, and the lead SNP in the region that was used in our MR analysis (rs174561), is located within the intron of host gene fatty acid desaturase 1 (*FADS1*) and 2 kb upstream of *FADS2*. Higher expression of miR-1908-5p has been associated with lower levels of plasma LDL-cholesterol (LDL-c), total cholesterol (TC), fasting glucose (FG), and glycated hemoglobin (HbA1c), and this effect is due to the regulatory impact of genetic variation in the region on circulating miR-1908-5p [16]. In this region, another CRC-associated locus (*TMEM258/MYRF*) is found; however, the mechanism behind its relationship with CRC is not clear [6, 51].

Genetically predicted miR-6810-3p was inversely associated with CRC risk; however, there is little evidence in the literature to support the role of miR-6810-3p in tumorigenesis. The variant that was used to proxy miR-6810-3p (rs1473901) is located in the region of *PNKD/TMBIM1*, a locus that has been previously associated with CRC [52]. Decreased cellular glutathione levels due to impaired PNKD function might increase oxidative stress levels, and TMBIM1 is implicated in modulating Fas ligand levels, both of which affect inflammation, a process linked to CRC initiation [52]. Enrichment

analyses showed that miR-6810-3p might be implicated in the downregulation of the ERBB4 signaling pathway, a pathway of emerging importance in CRC [53]. Among the associated proteins were WW domain containing E3 ubiquitin protein ligase 1 (*WWP1*), itchy E3 ubiquitin protein ligase (*ITCH*), and ubiquitin C (*UBC*), suggesting that miR-6810-3p potentially downregulates the ERBB4 pathway in a ubiquitination-dependent manner.

The potential of miRNAs as targets for cancer prevention and therapy has been actively investigated since their discovery in 1993 [54]. Research efforts include the development of miRNA inhibitors, which bind to miRNAs to block their function, and miRNA mimics, which imitate endogenous miRNAs' activity. Although most of the agents aimed at cancer therapy are currently in preclinical testing, a few, such as miR-16 and miR-34a mimics, have reached early-phase clinical trials [55–57]. However, challenges remain in characterizing their mechanisms thoroughly and addressing issues of sensitivity, specificity, selectivity, and off-target effects before clinical application is feasible. Moreover, lifestyle factors have been shown to modulate miRNA expression. For example, weight-loss interventions may alter extracellular levels of miR-146a-5p, potentially modulated further by physical activity [58, 59]. Experimental data also indicate that smoking upregulates miR-21, producing adverse effects in Caco-2 cell lines [60]. Additionally, phytochemicals, like curcumin and other phenolic compounds, have demonstrated anticancer activity in experimental models of hepatocellular and other cancers, partly through modulating miR-21 expression [61, 62].

Among the strengths of our analyses are the use of a wide range of biomarkers, including a comprehensive panel of miRNAs and proteins, exploring several mechanistic pathways behind CRC risk, using high-quality data. We used cis IVs, limiting potential pleiotropic effects, which was supported by the fact that pleiotropy scan showed little evidence revealed of pleiotropic pathways. There were only a few associations pertinent to metabolism and inflammation-related traits. However, considering the multilateral effects of miRNAs, this is likely a reflection of different mechanisms via which miRNAs might be associated with CRC risk, rather than horizontal pleiotropy.

Our study's primary limitation was that the GWAS of plasma miRNAs that we used in our main analysis (GWAS by Nikpay et al.) was restricted to individuals with obesity, and potential differences in the distribution of plasma miRNA concentrations compared to populations without obesity might have affected the genetic association estimates. When we compared the MR estimates with CRC risk using an alternative GWAS (by Huan et al.) with a proportion of individuals with obesity comparable to the

general population, the associations with CRC risk were qualitatively consistent providing some evidence of homogeneity in the genetic associations. It should be noted, however, that an additional source of variation in the estimates from the two GWAS is the difference in the samples used to quantify miRNAs, limiting comparability. In the GWAS by Nikpay et al., plasma samples were included, whereas the GWAS by Huan et al. included whole blood samples. Such a difference might explain why correlation was moderate. In a previous study, members of our team compared the genetic association estimates of the GWAS by Nikpay et al., with a GWAS in the Rotterdam Study (mean BMI in the study population of approximately 28 kg/m²) that used the same analytical platform [63, 64]. The majority of the associations were replicated, and the effect estimates of the replicated associations were strongly correlated ($r=0.82$), providing evidence to support homogeneity in the genetic associations across populations with marked differences in the prevalence of obesity [63].

Another important limitation is the use of a single SNP as IV for most of the analyses, which may have affected power to reject the null hypothesis for some associations and did not allow us to perform MR sensitivity analyses (i.e., weighted median, weighted mode, MR-Egger, and PRESSO). The sample size of the GWAS used to proxy plasma miRNA concentrations was relatively small; however, there was little evidence of weak instruments. Given that we used largely European populations to extract summary genetic association estimates for our analysis, generalizability to other populations is limited. In addition, there may be non-linear synergistic and time-dependent effects and biomarker-environment or biomarker-biomarker interactions that are not captured by the current analysis. In addition, parameters of gene expression, namely tissue specific and exposure specific expression, are not accounted for in MR analyses.

Conclusions

In conclusion, using high-throughput MR and colocalization analyses, we provide evidence that miR-146a-5p, miR-21-5p, miR-4707-3p, miR-1908-5p, and miR-6810-3p were associated with CRC risk. Additionally, several potential downstream protein targets and pertinent pathways are suggested, and their roles as intermediates in the miRNA to CRC associations should be further explored.

Abbreviations

CCFR	Colon Cancer Family Registry
CCL3	C–C motif chemokine ligand 3
CORECT	Colorectal Transdisciplinary Study
COX	Cyclooxygenase
CRC	Colorectal cancer
DAPK1	Death associated protein kinase 1
EMT	Epithelial mesenchymal transition

FADS1	Fatty acid desaturase 1
FAS	Fatty acid synthase
FDR	False discovery rate
FG	Fasting glucose
FHS	Framingham Heart Study
FUT9	Fucosyltransferase 9
GECCO	Genetics and Epidemiology of Colorectal Cancer Consortium
GL50	Growth inhibitory concentration
GSR	Glutathione-disulfide reductase
GWAS	Genome-wide association study
HbA1c	Glycated hemoglobin
IGF	Insulin-like growth factor
IGFBPs	Insulin-like growth factor binding proteins
IL1B	Interleukin 1B
ITCH	Itchy E3 ubiquitin protein ligase
IVs	Instrumental variables
IVW	Inverse variance weighting
KEGG	Kyoto Encyclopedia of Genes and Genomes
KL	Klotho
LD	Linkage disequilibrium
LDL-c	LDL-cholesterol
MAF	Minor allele frequency
MAPK8	Mitogen-activated protein kinase 8
MET	MET proto-oncogene, receptor tyrosine kinase
miR-eQTLs	MicroRNA expression quantitative trait loci
miRNAs	MicroRNAs
MR	Mendelian randomization
NCI	National Cancer Institute
NFKB1	Nuclear factor kappa B subunit 1
NRG1	Neuregulin 1
onco-miRs	Oncogenic miRNAs
PDGFRN	Prostaglandin F2 receptor inhibitor
PEA	Pathway enrichment analysis
PGE2	Prostaglandin E2
PID	Pathway Interaction Database
PP	Posterior probabilities
TC	Total cholesterol
TCGA	The Cancer Genome Atlas
TLR	Toll-like receptor
TP63	Tumor protein P63
ts-miRs	Tumor-suppressive miRNAs
UBC	Ubiquitin C
VCAM1	Vascular cell adhesion molecule 1
WebGestalt	WEB-based GEne Set Analysis Toolkit
WMe	Weighted median
Wmo	Weighted mode
WP	WikiPathways
WWP1	WW domain containing E3 ubiquitin protein ligase 1
ZNF750	Zinc finger protein 750
ZNRF3	Zinc and ring finger 3

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-025-04311-8>.

Additional file 1: Tables S1–S17. Table S1 Parameters for pathway enrichment analysis using the WEB-based GEne Set Analysis Toolkit. Table S2 Genetic association estimates of circulating miRNAs used in the Mendelian randomization analyses. Table S3 Mendelian randomization analyses of circulating miRNAs and risk of colorectal cancer using cis IVs. Table S4 Mendelian randomization analyses of circulating miRNAs and risk of colorectal cancer using trans IVs. Table S5 Summary of the colocalization analysis results. Table S6 Sensitivity Mendelian randomization analyses of circulating miRNAs and risk of colorectal cancer using cis IVs. Table S7 MR-link-2 analyses of miRNAs on colorectal cancer risk. Table S8 Pleiotropy scan of the cis IVs. Table S9 Mendelian randomization analysis of genetic liability to risk of developing colorectal cancer and plasma miRNA concentrations. Table S10 Mendelian randomization analyses of circulating miRNAs and plasma proteins. Table S11 Mendelian randomization analyses of

plasma proteins and colorectal cancer risk. Table S12 Suggestive miRNA and protein-target associations. Table S13 Mendelian randomization analyses plasma proteins on miRNAs. Table S14 Pathway enrichment analysis results. Table S15a miRNA expression profiles per tissue. Table S15b miRNA expression profiles per cancer tissue. Table S16 Correlation between miRNA and mRNA expression across tissues. Table S17 Correlation of miRNA expression across tissues and small compound sensitivity.

Additional file 2: Figures S1–S11. Fig. S1 Comparison of allele frequencies shared between the microRNA and protein GWAS. Fig. S2 Comparison of MR estimates for colorectal cancer risk using trans-defined instruments versus cis-defined instruments, regardless of significance. Fig. S3 Comparison of MR estimates for colorectal cancer risk using trans-defined instruments versus cis-defined instruments, focusing on significant cis-defined miRNA. Fig. S4 Regional plot of microRNA miR-146a-5p and colorectal cancer risk. Fig. S5 Regional plot of microRNA miR-21-5p and colorectal cancer risk. Fig. S6 Regional plot of micro-RNA miR-4707-3p and colorectal cancer risk. Fig. S7 Regional plot of microRNA miR-1908-5p and colorectal cancer risk. Fig. S8 Regional plot of microRNA miR-6810-3p and colorectal cancer risk. Fig. S9 Forest plot presenting the associations of the highlighted miRNAs with colorectal cancer subtypes, in Mendelian randomization inverse variance weighting analyses. Fig. S10 Comparative expression levels across healthy colon and blood tissues. Fig. S11 Correlation of expression levels between healthy colon tissue and plasma in log scale

Additional file 3: Funding and acknowledgements.

Acknowledgements

We are grateful to all the participants who have been part of the project and to the many members of the study teams. For more details see Additional file 3.

Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Authors' contributions

EB, CKP: Data curation, Formal analysis, Methodology, Investigation, Visualization, Writing - Original Draft, Writing - Review & Editing; RM, DS, SLS, AHW, HB, CIL, ATC, AJP, WZ, TOK, VM, CYU, BvG, AIP, RKP, SJL, RMM: Methodology, Writing - Review & Editing; AD, UP, MG: Methodology, Writing - Review & Editing, Supervision; KKT: Conceptualization, Funding acquisition, Investigation, Methodology, Writing - Review & Editing, Supervision. All authors critically revised the manuscript for intellectual content. All authors read and approved the final manuscript.

Funding

This work was supported by Cancer Research UK (grant number C18281/A29019). GECCO, CORECT, and CCFR funding and acknowledgements are described in detail (Additional file 3). RMM is a National Institute for Health Research Senior Investigator (NIHR202411). RMM is supported by a Cancer Research UK 25 (C18281/A29019) programme grant (the Integrative Cancer Epidemiology Programme). RMM is also supported by the NIHR Bristol Biomedical Research Centre which is funded by the NIHR (BRC-1215-20011) and is a partnership between University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. Department of Health and Social Care disclaimer: The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Data availability

All data used in this work are presented in the Additional files that accompany the manuscript and are described in the original publications. Full summary genetic association data for plasma miRNA concentrations can be found at <https://zenodo.org/records/2560974>, and for plasma proteins at <https://www.decode.com/summarydata/>. Researchers may have access to the summary-level genetic association data for colorectal cancer by submitting an application to GECCO.

Declarations

Ethics approval and consent to participate

All studies contributing summary statistics to these analyses had the relevant institutional review board approval from each country, in accordance with the Declaration of Helsinki. All participants provided informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, White City Campus, 90 Wood Lane, London W12 0BZ, UK. ²Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. ³Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK. ⁴Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH, USA. ⁵Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University School of Medicine, Cleveland, OH, USA. ⁶Preventative Medicine, University of Southern California, Los Angeles, CA, USA. ⁷Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. ⁸German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. ⁹Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA. ¹⁰Broad Institute of Harvard and MIT, Cambridge, MA, USA. ¹¹Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. ¹²Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ¹³Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. ¹⁴Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. ¹⁵Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ¹⁶Intermountain Health, Salt Lake City, UT, USA. ¹⁷Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, TN, USA. ¹⁸Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, NC, USA. ¹⁹Unit of Biomarkers and Susceptibility (UBS), Oncology Data Analytics Program (ODAP), Catalan Institute of Oncology (ICO), L'Hospitalet del Llobregat, Barcelona 08908, Spain. ²⁰Department of Population Science, American Cancer Society, Atlanta, GA, USA. ²¹Department of Diagnostics and Intervention, Oncology Unit, Umeå University, Umeå, Sweden. ²²Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden. ²³Department of Epidemiology, University of Washington, Seattle, WA, USA. ²⁴Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, AZ, USA. ²⁵Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. ²⁶NIHR Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK. ²⁷Nutrition and Metabolism Branch, International Agency for Research On Cancer, Lyon, France.

Received: 10 January 2025 Accepted: 31 July 2025

Published online: 02 October 2025

References

- Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–63.
- Jaspersion KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology*. 2010;138(6):2044–58.
- Jiao S, Peters U, Berndt S, Brenner H, Butterbach K, Caan BJ, Carlson CS, Chan AT, Chang-Claude J, Chanock S, et al. Estimating the heritability of colorectal cancer. *Hum Mol Genet*. 2014;23(14):3898–905.
- Nosho K, Kure S, Irahara N, Shima K, Baba Y, Spiegelman D, Meyerhardt JA, Giovannucci EL, Fuchs CS, Ogino S. A prospective cohort study shows unique epigenetic, genetic, and prognostic features of synchronous colorectal cancers. *Gastroenterology*. 2009;137(5):1609–1620.e1601–1603.
- Sun J, Zhao J, Jiang F, Wang L, Xiao Q, Han F, Chen J, Yuan S, Wei J, Larsson SC, et al. Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. *Genome Med*. 2023;15(1):75.
- Fernandez-Rozadilla C, Timofeeva M, Chen Z, Law P, Thomas M, Schmit S, Díez-Obrero V, Hsu L, Fernandez-Tajes J, Palles C, et al. Deciphering colorectal cancer genetics through multi-omic analysis of 100,204 cases and 154,587 controls of European and east Asian ancestries. *Nat Genet*. 2023;55(1):89–99.
- Vidman L, Zheng R, Bodén S, Ribbenstedt A, Gunter MJ, Palmqvist R, Harlid S, Brunius C, Van Guelpen B. Untargeted plasma metabolomics and risk of colorectal cancer—an analysis nested within a large-scale prospective cohort. *Cancer Metab*. 2023;11(1):17.
- Bartel DP. Metazoan microRNAs. *Cell*. 2018;173(1):20–51.
- Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019;47(D1):D155–d162.
- Bayraktar R, Van Roosbroeck K, Calin GA. Cell-to-cell communication: microRNAs as hormones. *Mol Oncol*. 2017;11(12):1673–86.
- Strubberg AM, Madison BB. MicroRNAs in the etiology of colorectal cancer: pathways and clinical implications. *Dis Model Mech*. 2017;10(3):197–214.
- Jung G, Hernández-Illán E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol*. 2020;17(2):111–30.
- Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res*. 2020;48(D1):D127–d131.
- Zhao C, Sun X, Li L. Biogenesis and function of extracellular miRNAs. *ExRNA*. 2019;1(1):38.
- Cojocneanu R, Braicu C, Raduly L, Jurj A, Zanoaga O, Magdo L, Irimie A, Muresan MS, Ionescu C, Grigorescu M, et al. Plasma and tissue specific miRNA expression pattern and functional analysis associated to colorectal cancer patients. *Cancers (Basel)*. 2020. <https://doi.org/10.3390/cancers12040843>.
- Nikpay M, Beehler K, Valsesia A, Hager J, Harper M-E, Dent R, McPherson R. Genome-wide identification of circulating miRNA expression quantitative trait loci reveals the role of several miRNAs in the regulation of cardiometabolic phenotypes. *Cardiovasc Res*. 2019;115(11):1629–45.
- Huan T, Rong J, Liu C, Zhang X, Tanriverdi K, Joehanes R, Chen BH, Murabito JM, Yao C, Courchesne P, et al. Genome-wide identification of microRNA expression quantitative trait loci. *Nat Commun*. 2015;6(1):6601.
- Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, Conti DV, Qu C, Jeon J, Edlund CK, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51(1):76–87.
- Ferkingsstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrnisdottir EL, Gunnarsdottir K, Helgason A, Oddsson A, Halldorsson BV, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53(12):1712–21.
- Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. 2011;40(3):755–64.
- Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658–65.
- Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103(4):965–78.
- van der Graaf A, Warmerdam R, Auwerx C, eQTLGen Consortium, Vösa U, Borges MC, Franke L, Kutalik Z. MR-link-2: pleiotropy robust cis Mendelian randomization validated in three independent reference datasets of causality. *Nat Commun*. 2025;16(1):6112. <https://doi.org/10.1038/s41467-025-60868-1>. PMID: 40610416; PMCID: PMC12229666.
- Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS, Staley JR. Phenoscanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics (Oxford, England)*. 2019;35(22):4851–3.
- Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, et al. Phenoscanner: a database of human genotype-phenotype associations. *Bioinformatics (Oxford, England)*. 2016;32(20):3207–9.
- Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–14.

27. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32(5):377–89.
28. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–25.
29. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46(6):1985–98.
30. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383.
31. Karhunen V, Gill D, Huang J, Bouras E, Malik R, Ponsford MJ, Ahola-Olli A, Papadopoulou A, Palaniswamy S, Sebert S, et al. The interplay between inflammatory cytokines and cardiometabolic disease: bi-directional mendelian randomisation study. *BMJ Med*. 2023;2(1):e000157.
32. Liao Y, Wang J, Jaehnnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47(W1):W199–w205.
33. Rishik S, Hirsch P, Grandke F, Fehlmann T, Keller A. miRNATissueAtlas 2025: an update to the uniformly processed and annotated human and mouse non-coding RNA tissue atlas. *Nucleic Acids Res*. 2025;53(D1):D129–d137.
34. Liu CJ, Fu X, Xia M, Zhang Q, Gu Z, Guo AY. Mirnasnp-v3: a comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. *Nucleic Acids Res*. 2021;49(D1):D1276–d1281.
35. Luna A, Elloumi F, Varma S, Wang Y, Rajapakse VN, Aladjem MI, Robert J, Sander C, Pommier Y, Reinhold WC. Cell miner cross-database (CellMiner-CDB) version 1.2: exploration of patient-derived cancer cell line pharmacogenomics. *Nucleic Acids Res*. 2021;49(D1):D1083–93.
36. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. URL <https://www.R-project.org/>.
37. Slattery ML, Mullany LE, Sakoda L, Samowitz WS, Wolff RK, Stevens JR, Herrick JS. The NF- κ B signalling pathway in colorectal cancer: associations between dysregulated gene and miRNA expression. *J Cancer Res Clin Oncol*. 2018;144(2):269–83.
38. Hwang WL, Jiang JK, Yang SH, Huang TS, Lan HY, Teng HW, Yang CY, Tsai YP, Lin CH, Wang HW, et al. MicroRNA-146a directs the symmetric division of Snail-dominant colorectal cancer stem cells. *Nat Cell Biol*. 2014;16(3):268–80.
39. Boldin MP, Taganov KD, Rao DS, Yang L, Zhao JL, Kalvani M, Garcia-Flores Y, Luong M, Devrekanli A, Xu J, et al. Mir-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *J Exp Med*. 2011;208(6):1189–201.
40. Garo LP, Ajay AK, Fujiwara M, Gabriely G, Raheja R, Kuhn C, Kenyon B, Skillin N, Kadowaki-Saga R, Saxena S, et al. MicroRNA-146a limits tumorigenic inflammation in colorectal cancer. *Nat Commun*. 2021;12(1):2419.
41. Saba R, Sorensen DL, Booth SA. MicroRNA-146a: a dominant, negative regulator of the innate immune response. *Front Immunol*. 2014;5:578.
42. Orsini M, Chateauvieux S, Rhim J, Gaigneaux A, Cheillan D, Christov C, Dicato M, Morceau F, Diederich M. Sphingolipid-mediated inflammatory signaling leading to autophagy inhibition converts erythropoiesis to myelopoiesis in human hematopoietic stem/progenitor cells. *Cell Death Differ*. 2019;26(9):1796–812.
43. Wu Y, Song Y, Xiong Y, Wang X, Xu K, Han B, Bai Y, Li L, Zhang Y, Zhou L. MicroRNA-21 (Mir-21) promotes cell growth and invasion by repressing tumor suppressor PTEN in colorectal cancer. *Cell Physiol Biochem*. 2017;43(3):945–58.
44. Slattery ML, Mullany LE, Sakoda LC, Samowitz WS, Wolff RK, Stevens JR, Herrick JS. Expression of Wnt-signaling pathway genes and their associations with miRNAs in colorectal cancer. *Oncotarget*. 2018;9(5):6075–85.
45. Peacock O, Lee AC, Cameron F, Tarbox R, Vafadar-Isfahani N, Tufarelli C, Lund JN. Inflammation and miR-21 pathways functionally interact to downregulate PDCC4 in colorectal cancer. *PLoS ONE*. 2014;9(10):e110267.
46. Titz-de-Almeida R, David C, Titz-de-Almeida SS. The race of 10 synthetic RNAi-based drugs to the pharmaceutical market. *Pharm Res*. 2017;34(7):1339–63.
47. Bi Y, Guo S, Xu X, Kong P, Cui H, Yan T, Ma Y, Cheng Y, Chen Y, Liu X, et al. Decreased ZNF750 promotes angiogenesis in a paracrine manner via activating DANCR/miR-4707-3p/FOXC2 axis in esophageal squamous cell carcinoma. *Cell Death Dis*. 2020;11(4):296.
48. Liu Y, Zhang M, Liang L, Li J, Chen YX. Over-expression of lncRNA DANCR is associated with advanced tumor progression and poor prognosis in patients with colorectal cancer. *Int J Clin Exp Pathol*. 2015;8(9):11480–4.
49. Cui YM, Jiao HL, Ye YP, Chen CM, Wang JX, Tang N, Li TT, Lin J, Qi L, Wu P, et al. FOXC2 promotes colorectal cancer metastasis by directly targeting MET. *Oncogene*. 2015;34(33):4379–90.
50. Shen J, Wu Y, Ruan W, Zhu F, Duan S. Mir-1908 dysregulation in human cancers. *Front Oncol*. 2022;12:857743.
51. Zhang B, Jia WH, Matsuda K, Kweon SS, Matsuo K, Xiang YB, Shin A, Jee SH, Kim DH, Cai Q, et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat Genet*. 2014;46(6):533–42.
52. Orlando G, Law PJ, Palin K, Tuupanen S, Gylfe A, Hänninen UA, Cajuso T, Tanskanen T, Kondelin J, Kaasinen E, et al. Variation at 2q35 (PNKD and TMBIM1) influences colorectal cancer risk and identifies a pleiotropic effect with inflammatory bowel disease. *Hum Mol Genet*. 2016;25(11):2349–59.
53. Segers VFM, Dugaucquier L, Feyen E, Shakeri H, De Keulenaer GW. The role of ErbB4 in cancer. *Cell Oncol*. 2020;43(3):335–52.
54. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*. 1993;75(5):843–54.
55. Seyhan AA. Trials and tribulations of microRNA therapeutics. *Int J Mol Sci*. 2024. <https://doi.org/10.3390/ijms25031469>.
56. van Zandwijk N, Pavlakis N, Kao SC, Linton A, Boyer MJ, Clarke S, Huynh Y, Chrzanowska A, Fulham MJ, Bailey DL, et al. Safety and activity of microRNA-loaded minicells in patients with recurrent malignant pleural mesothelioma: a first-in-man, phase 1, open-label, dose-escalation study. *Lancet Oncol*. 2017;18(10):1386–96.
57. Hong DS, Kang YK, Borad M, Sachdev J, Ejadi S, Lim HY, Brenner AJ, Park K, Lee JL, Kim TY, et al. Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *Br J Cancer*. 2020;122(11):1630–7.
58. Veie CHB, Nielsen IMT, Frisk NLS, Dalgaard LT. Extracellular microRNAs in relation to weight loss—a systematic review and meta-analysis. *Non-coding RNA*. 2023;9(5):53.
59. Russo A, Bartolini D, Mensà E, Torquato P, Albertini MC, Olivieri F, Testa R, Rossi S, Piroddi M, Cruciani G, et al. Physical activity modulates the overexpression of the inflammatory miR-146a-5p in obese patients. *IUBMB Life*. 2018;70(10):1012–22.
60. Dino P, D’Anna C, Sangiorgi C, Di Sano C, Di Vincenzo S, Ferraro M, Pace E. Cigarette smoke extract modulates E-cadherin, claudin-1 and miR-21 and promotes cancer invasiveness in human colorectal adenocarcinoma cells. *Toxicol Lett*. 2019;317:102–9.
61. Li J, Wei H, Liu Y, Li Q, Guo H, Guo Y, Chang Z. Curcumin inhibits hepatocellular carcinoma via regulating miR-21/TIMP3 axis. *Evid Based Complement Altern Med*. 2020;2020:2892917.
62. Hashemi M, Mirdamadi MSA, Talebi Y, Khaniabad N, Banaei G, Daneii P, Gholami S, Ghorbani A, Tavakolpournegari A, Farsani ZM, et al. Pre-clinical and clinical importance of miR-21 in human cancers: tumorigenesis, therapy response, delivery approaches and targeting agents. *Pharmacol Res*. 2023;187:106568.
63. Mustafa R, Mens MMJ, van Hilten A, Huang J, Roshchupkin G, Huan T, Broer L, van Meurs JBJ, Elliott P, Levy D, Ikram MA, Evangelou M, Dehghan A, Ghanbari M. A comprehensive study of genetic regulation and disease associations of plasma circulatory microRNAs using population-level data. *Genome Biol*. 2024;25(1):276. <https://doi.org/10.1186/s13059-024-03420-6>. PMID: 39434104; PMCID: PMC11492503.
64. Ikram MA, Brussels G, Ghanbari M, Goedegebuure A, Ikram MK, Kavousi M, Kieboom BCT, Klaver CCW, de Knecht RJ, Luik AI, et al. Objectives, design and main findings until 2020 from the Rotterdam study. *Eur J Epidemiol*. 2020;35(5):483–517.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.