1 **Sulforaphane modulates microRNA expression in colon cancer cells to**
 2 **implicate the regulation of oncogenes CDC25A, HMGA2 and MYC.**

3 Christopher A. Dacosta¹, Claudia Paicu^{2,3}, Irina Mohorianu^{2,3}, Wei Wang¹, Ping Xu³, Tamas
 4 Dalmay³, Yongping Bao^{1*}

5 ¹Norwich Medical School, University of East Anglia, Norwich, NR4 7UQ, UK

6 ²School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

7 ³School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

8 *Requests for reprints: Yongping Bao, Norwich Medical School, University of East Anglia,
 9 Norwich, NR4 7UQ, UK. Tel: +44 1603 591778; E-mail: y.bao@uea.ac.uk

10 **Running Title:** SFN against colorectal cancer: potential roles of miRNAs.

11 **Keywords:** sulforaphane; microRNA (miRNA); cancer; colorectal; cruciferous vegetable.

12 **Conflicts of Interest:** There are no conflicts of interest for any of the authors.

13

14 **Abstract**

15 Colorectal cancer is an increasingly important cause of morbidity and mortality, whose
16 incidence is associated with dietary and lifestyle factors, particularly inversely so with the
17 consumption of cruciferous vegetables. These vegetables contain glucosinolates, from the
18 breakdown of which are derived isothiocyanates, such as sulforaphane. Sulforaphane is well-
19 characterised for wide-ranging tumour-suppressive and chemoprotective activities *in vitro*,
20 yet deeper elucidation of its biological interactions would aid in better realising its potential
21 in chemoprevention and/or chemotherapy. There is evidence to suggest that sulforaphane
22 modulates microRNA expression in the colon, thus implying the potential for microRNA
23 modulation to play a role in the anti-cancer effects of sulforaphane. Therefore, the effects of
24 sulforaphane on microRNA expression profiles in the colonic adenocarcinoma Caco-2 and
25 non-cancerous colonic CCD-841 cell lines were investigated by small RNA cloning and deep
26 sequencing, followed by Northern Blot validation experiments. Sulforaphane upregulated let-
27 7f-5p and let-7g-5p expression at 24 h in Caco-2 cells, but not in CCD-841. Such treatment
28 also downregulated miR-29b-3p in Caco-2. Dual luciferase assays with a let-7f-5p mimic and
29 inhibitor confirmed the binding of the miRNA to predicted binding sites in the mRNA
30 transcript 3'-UTRs of cell division cycle 25A (CDC25A), high-mobility group AT-hook-2
31 (HMGA2) and MYC. Therefore, we hypothesize that let-7f-5p translationally represses
32 CDC25A, HMGA2 and MYC, thereby playing a role in the tumour-suppressive effects of
33 sulforaphane. The apparent selectivity of let-7f-5p induction towards tumour cells would be
34 therapeutically desirable if applicable *in vivo*. MiR-29b-3p is predicted to target a number of
35 tumour-suppressing genes, further investigation of which could be informative regarding the
36 potential of sulforaphane to suppress tumour progression.

37

38 Introduction

39 Colorectal cancer is an increasingly major cause of global morbidity and mortality, the latter
 40 of which increased by 46% between 1990 and 2010 (1). Its global incidence was 1.4 million
 41 in 2012 (2), making it the third most commonly diagnosed type of cancer (3). It originates
 42 from the oncogenic transformation of epithelial cells lining the colon or rectum, which are
 43 particularly susceptible to the oxidative, mutagenic and/or inflammatory effects of ingested
 44 compounds or products of the gut microbiota. Therefore, it is unsurprising that its incidence
 45 has been repeatedly linked to lifestyle and dietary factors, and observed to rise with the
 46 adoption of ‘Western’ dietary habits in a given country (4). Preventative strategies and the
 47 screening of asymptomatic individuals for early-stage colorectal cancer are vital, because
 48 noticeable symptoms tend not to arise until the cancer has progressed to an advanced and
 49 often terminal stage (5).

50 Colorectal cancer risk is particularly inversely correlated with cruciferous vegetable
 51 consumption (6), which is likely to be attributable to the derivation of isothiocyanates (ITCs)
 52 from the glucosinolates: a group of compounds particular to this group of vegetables that
 53 includes broccoli, Brussel sprouts, cabbage and bok choy. Plant cell rupture (e.g. by cooking
 54 and/or chewing) enables myrosinase enzymes to hydrolyse the glucosinolates to form several
 55 types of product, including ITCs. Broccoli is high in a particular glucosinolate,
 56 glucoraphanin, from which is derived a specific ITC known as sulforaphane (SFN). SFN,
 57 amongst other ITCs, is a well-established inducer of the nuclear factor (erythroid-derived 2)-
 58 like 2 (Nrf2) transcription factor, which upregulates the transcription of various antioxidant
 59 genes whose promoters contain the *antioxidant response element*. These help to protect
 60 against potentially oncogenic, oxidation-induced, mutagenesis, such as by inhibiting 2-
 61 amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP)-induced DNA damage (7). It is also

62 widely reported to directly suppress oncogenic characteristics of tumour cell lines, such as
63 cell migration, invasion and proliferation, and to promote cell cycle arrest and apoptosis in
64 such (8).

65 Moreover, SFN is a hormetic agent, in that the medium-term consequences of low-to-
66 moderate exposure – i.e. that reasonably achievable in humans via vegetable consumption -
67 tend to be cytoprotective, whereas much higher-dose exposure can be cytotoxic via the
68 induction of oxidative stress and DNA damage (9). This is because SFN acts acutely as a pro-
69 oxidant via several mechanisms, including the depletion of cellular glutathione (10) and the
70 disruption of the mitochondrial respiratory chain (11). Low-to-moderate SFN exposure can
71 trigger an Nrf2-mediated anti-oxidative cellular response whose reductive effects outweigh
72 those of the initial pro-oxidant effect in the medium term. On the other hand, the pro-oxidant
73 effects of high-dose exposure may be such as to induce significant oxidative damage to
74 macromolecules and thus result in cytotoxicity rather than cytoprotection. The latter form of
75 activity is obviously undesirable if occurring in healthy cells, but inevitably contributes to the
76 tumour-suppressive effects of SFN observed in tumour cell lines, as described above.
77 Conversely, while the cytoprotective capacity of low-to-moderate dose SFN is desirable in
78 normal cells and helps to protect them against oncogenic transformation, such cytoprotection
79 of tumour cells is undesirable as it may increase their resistance to cytotoxicity-dependent
80 forms of anti-cancer therapy. This has led to speculation that low-to-moderate doses of SFN,
81 whether dietary or supplemental, might even be detrimental to patients undergoing treatment
82 for advanced cancers. However, the ‘cytotoxic threshold’ of SFN with respect to dose
83 inevitably varies according to cell type, due to variations in basal levels of reactive oxygen
84 species and cell signalling. Notably, a number of studies have indicated tumour cell lines to
85 be more sensitive than their non-cancerous counterparts to SFN-induced cytotoxicity (i.e. to
86 have a lower ‘cytotoxic threshold’) (12), which opens up the possibility of using SFN at

87 intermediate doses that are simultaneously cytotoxic towards tumour cells yet cytoprotective
88 towards surrounding non-cancerous cells.

89 Clearly, the bioactivity of SFN is highly complex and far from fully understood. Further
90 understanding of SFN's biological interactions is crucial in order to further investigate its
91 potential chemopreventative and/or chemotherapeutic uses. There is evidence to suggest that
92 SFN is able to modulate the expression of miRNAs in several colonic cell lines, from a study
93 by Slaby et al. (13). Given that miRNAs are well-established as major players in human
94 development and disease progression via their post-transcriptional regulation of gene
95 expression, it is highly probable that SFN's anti-cancer effects are partly attributable to the
96 modulation of miRNA expression. Therefore, a study to investigate the modulation of
97 miRNA expression by SFN in a colonic adenocarcinoma cell line, Caco-2, and a non-
98 cancerous colonic cell line, CCD-841, was carried out by first performing a wide-scale assay
99 of SFN-mediated miRNA modulation by small RNA (sRNA) library cloning and deep
100 sequencing, and then verifying the SFN-mediated modulation of specific miRNAs by
101 individual assays. Subsequently, the ability of modulated miRNAs to bind to several
102 computationally-predicted oncogenic mRNA targets was verified *in vitro*. The insight into the
103 potential bioactivity of SFN provided by the findings of the study with respect to colorectal
104 carcinogenesis could lead to new avenues of research regarding the potential applications of
105 diet-derived and/or supplemental SFN in chemoprevention, and/or standalone or adjuvant
106 chemotherapy.

Materials and Methods

Materials

SFN (1-isothiocyanato-4-[methylsulfinyl]-butane, 99% purity (Cat No. S699115) was purchased from Toronto Research Chemicals (Canada). Dimethyl sulfoxide (DMSO), Dulbecco's Modified Eagle's Medium containing D-glucose (4.5g/L) and non-essential amino acids (DMEM) (Cat. no. D5671-500ML), heat-inactivated FBS (Cat. no. 12106C-500ML), L-glutamine (Cat. no. 59202C-100ML), Penicillin-Streptomycin (Cat. no. P4083-100ML), DNA oligonucleotides, TRI Reagent (Cat. no. T9424), the mirPremier™ MicroRNA Isolation Kit (Cat. no. SNC50), and the PerfectHyb™ Plus Hybridization Buffer (Cat no. H7033) were all purchased from Sigma-Aldrich (Dorset, UK). The let-7f-5p mimic (Cat. no. MSY0000067; target sequence: 5'-UGAGGUAGUAGAUUGUAUAGUU-3'), let-7f-5p inhibitor (Cat. no. MIN0000067; target sequence: 5'-UGAGGUAGUAGAUUGUAUAGUU-3'), AllStars Negative Control siRNA (Cat. no. SI03650318), and Attractene Transfection Reagent (Cat. no. 301005) were all purchased from Qiagen (West Sussex, UK). RNase Inhibitor (Cat. no. AM2682), Gel Loading Buffer II (Cat no. AM8546G), T4 Polynucleotide Kinase Forward Reaction Buffer (Cat. no. 10343010) and T4 Polynucleotide Kinase (Cat. no. EK0032) were all purchased from Fisher Scientific (Loughborough, UK). The Zymo RNA Clean & Concentrator™ -25 kit (Cat. no. R1017) was purchased from Cambridge Bioscience (Cambridge, UK). The Dual-Glo® Luciferase Assay System (Cat. no. E2920) was purchased from Promega UK (Southampton, UK). The human colorectal adenocarcinoma Caco-2 (HTB-37) and human normal colonic CCD-841 (CRL-1790) cell lines were obtained from ATCC.

129 Cell Culture and Treatments

130 Cells were maintained in DMEM supplemented with FBS (10%) (v/v), L-glutamine
 131 (200nM), and Penicillin-Streptomycin. Cells were cultured at 37°C in air under 5% CO₂.
 132 Caco-2 cells were split before exceeding 80% confluence, to avoid their differentiation.
 133 Caco-2 cells were tested in the laboratory and found to be mycoplasma free. CCD-841 cells
 134 were not tested for mycoplasma in the laboratory.

135 SFN was dissolved in DMSO, and final DMSO concentrations were kept at 0.05%. For
 136 sRNA library construction and Northern Blotting experiments, Caco-2 cells were cultured to
 137 70-80% confluence in 10cm dishes, then treated in duplicate with SFN (10μM) in culture
 138 medium for 8 or 24 h, or DMSO (control). For luciferase assays, HCT-116 cells were subject
 139 to transfection with pmiRGLO vectors and miRNA mimic and/or inhibitor at the same time
 140 as the seeding of 3.0×10^4 cells/well in 96-well plates. Co-transfection of let-7f-5p mimic
 141 and/or let-7f-5p inhibitor and/or AllStars Negative Control siRNA was performed using
 142 Attractene Transfection Reagent, according to the manufacturer's instructions. Briefly,
 143 transfection mixtures were prepared by combining Attractene, pmiRGLO vector, and miRNA
 144 mimic (100nM) and/or inhibitor (100nM) and/or AllStars Negative Control siRNA (0-
 145 200nM), in non-supplemented DMEM, adjusting the amounts of AllStars Negative Control
 146 siRNA in order to maintain a 200nM RNA concentration. These mixtures were incubated at
 147 room temperature for 15 min, and then added directly into suspended cells immediately
 148 following the seeding of cells into 96-well plates. Cells were subsequently incubated for 48 h.

RNA Extraction

For sRNA library construction, total RNA was extracted using TRI Reagent, according to the manufacturer's instructions (with minor modifications to the protocol). Briefly, cells were lysed in TRI Reagent by directly adding reagent onto attached cells for ≥ 1 min following the removal of culture medium. Lysates were incubated for 5 min at room temperature, then 200 μ L chloroform was mixed per 1mL lysate. Mixtures were incubated at room temperature for 10 min after vigorous shaking, then centrifuged at 16 000 g for 15 min at 4°C. The resulting aqueous, uppermost phases were retained, from which RNA was then recovered by isopropanol precipitation (≥ 5 h at -20°C) and then centrifugation at ≥ 16 000 g for 30 min. Pellets were recovered, washed in 75% ethanol, dried, and then resuspended in nuclease-free H₂O with RNase Inhibitor.

For Northern Blot experiments, either total RNA was extracted in the same manner as that described above, or small RNA was extracted using the mirPremier™ MicroRNA Isolation Kit according to the manufacturer's instructions.

Small RNA Library Construction

Total RNA samples were integrity-checked by agarose gel electrophoresis to visualise 28S and 18S bands, and then purified using the Zymo RNA Clean & Concentrator™ -25 kit according to the manufacturer's instructions. Subsequently, sRNA libraries were constructed from the purified total RNA samples according to the protocol published in 2015 by Xu et al. (14) with the HD adapters previously developed by Sorefan et al. (15), using 2 μ g total RNA per library.

Small RNA Library Sequencing and Bioinformatics Analysis

Nine indexed sRNA libraries were multiplexed on a sequencing lane and deep-sequenced on the Illumina HiSeq 2500 platform by the Earlham Institute, formerly known as The Genome Analysis Centre (Norwich, UK), with a sequencing depth of 100 million reads per lane.

Analysis of the deep sequencing data was performed using custom-made Perl and R scripts. Firstly, the adapters were trimmed based on a perfect match to the first 8 nt of the adapter sequence (TGGAATTC). Next, the HD signatures (assigned *NNNN* fragments localized at the 5' and 3' ends of each read) were trimmed. The reads were converted into non-redundant format. The unique sequences and their abundances i.e. the number of occurrences in a sample, were logged. The reads were then mapped to the reference human genome (*H. sapiens* v. 38 (16)), the available annotations, and the *H. sapiens* miRNAs downloaded from miRBase v. 21 (17), using PatMaN (18), allowing 0 gaps and ≤ 2 mismatches (see Supplementary Table S1 for a summary of raw and accepted read numbers per library). Next, the read abundances were normalized by the reads per million (RPM) (19), sRNA-adapted quantile normalization (20) and subsampling (21) methods. The most appropriate normalization method for this dataset was determined to be RPM, conducted on a scaling total of 13 000 000: the median total value. The differential expression analysis was done on the median expression levels of the biological replicates as pairwise comparisons, using an offset fold-change approach (offset=20) (22). All miRNAs with a fold-change ≥ 1.9 between either control and 8 h, control and 24 h, or 8 h and 24 h, were predicted as differentially-expressed candidates.

The data corresponding to this study are publicly available in the Gene Expression Omnibus (GEO) (23) under accession number GSE89363 (samples GSM2367444 to GSM2367459).

For library size class distributions, see Supplementary Figure S1. For complexity distributions, see Supplementary Figure S2. For the read annotation analysis and determination of differential expression, see Supplementary Tables S2, S3 and S4.

Northern Blotting

Equal amounts of total RNA (10µg) or small RNA (5µg) were combined with Gel Loading Buffer II, denatured at 65°C for 5 min, then size-fractionated by urea-PAGE in pre-run 16% gels at 100V for 2 h, in 0.5X tris/borate/EDTA buffer (TBE). Size-fractionated RNA was transferred to Hybond® NX membranes by semi-dry electroblotting in 0.5X TBE at 3mA/cm² for 1 h. RNA was then chemically cross-linked to membranes using a solution prepared by diluting 0.373g 1-ethyl-3-(dimethylaminopropyl) carbodiimide in 10mL dH₂O, then adding 122.5µL 12.5M 1-methylimidazole and 10µL 1M HCl. Membranes were placed face-up onto filter paper soaked with the solution, sealed in Saran wrap, then incubated at 60°C for 1 – 2 h. Following cross-linking, membranes were washed in dH₂O for 10 min.

PerfectHyb™ Plus hybridization buffer was pre-incubated at 37°C to dissolve precipitates, then membranes were incubated in hybridization buffer at 37°C under rotation for 1 h. Radioactive miRNA probes were prepared by mixing T4 Polynucleotide Kinase Forward Reaction Buffer, DNA oligonucleotide antisense to the miRNA of interest (or part of human U6) (10µM), T4 Polynucleotide Kinase, dH₂O, and γ-ATP (P-32), then incubating at 37°C for 1 h. Prepared radioactive probes were added to hybridization buffer, in which membranes were subsequently incubated at 37°C under rotation overnight. Then, membranes were washed four times in 0.2X SSC/0.1% SDS solution and sealed in Saran wrap. Phosphorimaging plates were exposed to the radioactively-labelled membranes for appropriate periods of time (typically 6-24 h), and then signals were evaluated using a

phosphorimaging scanner. RNA band intensities were densitometrically analyzed using the ImageJ software. Prior to re-probing for different targets, membranes were stripped by incubation in near-boiling 0.1% SDS.

Dot Blotting

Dot blots were prepared by adding DNA oligonucleotides of sequence sense to those of miRNAs of interest directly onto Hybond® NX membranes, then chemically cross-linking them to the membranes in the same manner that blotted RNA was cross-linked to membranes for Northern Blotting. Radioactive probe preparation, hybridization, and phosphorimaging were subsequently performed in the same manner as for Northern Blotting.

Luciferase Reporter Assays

Dual luciferase reporter assays (Firefly/*Renilla*) were performed using the Dual-Glo® Luciferase Assay System, according to the manufacturer's instructions, and luminescence was measured using the FLUOstar Omega microplate reader (BMG Labtech), using non-transfected cells to determine background luminescence. Firefly luminescence values were normalized to *Renilla* luminescence values to control for any well-to-well variation in cell viability and/or transfection efficiency.

Results

Sulforaphane-Induced MicroRNA Differential Expression

The sRNA-seq data analysis indicated that the treatment of Caco-2 cells with SFN (10μM) for 8 or 24 h modulated the expression of 42 miRNAs by ≥ 1.9 -fold, based on pairwise

comparisons between time points of the medians of replicate normalized expression levels. Overall, let-7f-5p and let-7g-5p were the most abundant mature miRNAs, and were upregulated 1.4- and 2.1-fold by 8 and 24 h SFN treatment, respectively, as illustrated by the data presented in Figure 1. Reads for miR-193b-3p were of moderate abundance, and indicated 1.5- and 2.0-fold downregulation by 8 and 24 h SFN treatment, respectively, as illustrated by the data presented in Figure 1.

Northern Blot experiments confirmed that SFN (10 μ M) upregulated let-7f-5p and let-7g-5p expression in Caco-2 cells; the former by 1.2- and 1.4-fold at 8 and 24 h respectively, and the latter by 1.1- and 1.6-fold at 8 and 24 h respectively, as illustrated by the data presented in Figure 2 [A]. However, they showed that similar treatment of CCD-841 cells affected the expression of neither let-7f-5p nor let-7g-5p, as illustrated by the data presented in Figure 2 [B]. Meanwhile, Northern Blots for miR-193b-3p confirmed a downregulation in Caco-2 by 8 but not 24 h SFN treatment, as illustrated by the data presented in Figure 3 [A]. The Northern Blot experiment for miR-29b-3p indicated that it was downregulated in Caco-2 cells 1.5- and 1.7-fold by 8 and 24 h SFN treatment, respectively, as illustrated by the data presented in Figure 3 [B].

Specificity of Let-7f-5p and Let-7g-5p Northern Blot Probe Hybridization

Many miRNAs of the let-7 family are very similar in sequence, varying only at loci outside of the seed region. For example, let-7f-5p and let-7g-5p vary at positions 12 and 18, where let-7f-5p has adenine and uracil, while let-7g-5p has uracil and cytosine. This led to initial speculation that the let-7f-5p and/or let-7g-5p probe(s) used for Northern Blotting might have bound off-target to other let-7 miRNAs. In order to test the potential for such cross-

hybridization, dot-blots of DNA oligonucleotides sense to the sequences of let-7a-5p, let-7b-5p, let-7c-5p, let-7d-5p, let-7e-5p, let-7f-5p, let-7g-5p and let-7i-5p were prepared, and then probed using the let-7f-5p and let-7g-5p Northern Blot probes.

Densitometric dot analysis indicated that the let-7f-5p probe was largely specific to let-7f-5p, albeit binding to let-7a-5p with affinity 6.3-fold lower than that with which it bound its target, as illustrated by the data presented in Figure 4, which also demonstrate that the let-7g-5p Northern Blot probe was largely specific for its target, albeit binding to let-7i-5p with 5.3-fold lower affinity.

However, reads for neither let-7a-5p nor let-7i-5p were actually present in the Caco-2 sRNA library deep sequencing data.

Targeting of mRNA 3-Untranslated Regions by Let-7f-5p

Targets of let-7f-5p predicted by the miRanda algorithm with mirSVR scores ≤ -1.2 were obtained from the online *microRNA.org – Targets and Expression* database (24), and included the mRNA transcripts of CDC25A, HMGA2, MYC and KRAS.

Let-7f-5p Antisense-Repeats

In the luciferase assay experiments, co-transfection of the let-7f-5p mimic with the let-7f-5p antisense repeat-containing vector repressed relative Firefly luciferase expression 3.2-fold, while co-transfection of the inhibitor with the same vector upregulated relative Firefly luciferase expression 3.0-fold, as illustrated the data presented in Figure 5.

Co-transfection of both the mimic and inhibitor with the vector resulted in no net modulation of relative Firefly luciferase expression.

CDC25A, HMGA2, MYC and KRAS mRNA 3'-Untranslated Regions

Co-transfection of the let-7f-5p mimic with the vector containing part of the CDC25A 3'-UTR with a computationally-predicted let-7f-5p binding site did not affect relative Firefly luciferase expression. However, co-transfection of the let-7f-5p inhibitor with the same vector upregulated relative Firefly luciferase expression 1.4-fold, as the data presented in Figure 6 [A] demonstrate. Co-transfection of either the mimic or inhibitor with the vector containing a modified form of the CDC25A 3'-UTR portion, in which the predicted binding site was mutated, had no effect.

The co-transfection of the mimic with the vector containing part of the HMGA2 3'-UTR, with two predicted binding sites, repressed relative Firefly luciferase expression 1.6-fold, as the data presented in Figure 6 [B] illustrate. However, co-transfection of the inhibitor with the vector had no significant effect, but the co-transfection of inhibitor together with the mimic and vector clearly abrogated the mimic's repressive effects. Neither the mimic nor inhibitor affected relative Firefly luciferase expression when co-transfected with the mutated-binding-site form of the vector.

Co-transfection of the mimic with the vector containing part of the MYC 3'-UTR had no effect on relative Firefly luciferase expression, but co-transfection of the inhibitor increased such expression 1.4-fold, as illustrated by the data presented in Figure 6 [C]. Neither the mimic nor inhibitor had any effect on relative Firefly luciferase expression when solely co-

transfected with the mutant form of the vector, but when the mimic, inhibitor and vector were all co-transfected at once, expression from the mutant vector unexpectedly increased 2.2-fold.

The co-transfection of neither the mimic nor inhibitor had any significant effect on relative Firefly luciferase expression from the vector containing part of the KRAS 3'-UTR with one predicted binding site, as illustrated by the data presented in Figure 6 [D]. However, co-transfection of the mimic with the mutant-UTR vector unexpectedly increased relative Firefly expression 1.6-fold, while transfection of the inhibitor together with the mimic and vector partially abrogated the mimic's effects.

Discussion

The miRNA expression profiling data show that SFN (10 μ M) is able to upregulate the expression of both let-7f-5p and let-7g-5p in the colonic adenocarcinoma Caco-2 cells, but not in the non-cancerous colonic epithelial CCD-841 cells. The mechanism behind this differential effect between the cell lines is unclear. One possibility is that SFN treatment suppresses NF- κ B in Caco-2 cells, thereby inhibiting lin28 and thus upregulating let-7 expression (25). The luciferase assay data demonstrate that let-7f-5p is indeed able to bind directly to a locus in the CDC25A 3'-UTR that was computationally predicted as a let-7f-5p binding site, and that said locus would play a role in the potential let-7f-5p-mediated translational repression of CDC25A. They also confirm that the same miRNA is able to directly bind one or more loci in the HMGA2 3'-UTR that were predicted as binding sites, and that said loci could play a role in the potential let-7f-5p-mediated translational repression of HMGA2. Finally, they suggest that the miRNA is able to directly bind to a computationally-predicted binding site in the 3'-UTR of MYC, and that this site would play a role in potential let-7f-5p-induced repression of MYC translation.

Overall, the data confirm that let-7f-5p and let-7g-5p were upregulated in Caco-2, but not in CCD-841, by SFN (10 μ M) treatment, and that let-7f-5p can directly interact with loci in the oncogenic CDC25A, HMGA2 and MYC 3'-UTRs. The let-7 family of miRNAs are widely reported to have anti-proliferative, apoptotic and/or cell cycle arrest-inducing effects across a number of cell types. Thus, the data presented here indicate that the translational repression of CDC25A, HMGA2 and MYC is likely to contribute to the colorectal cancer-suppressive effects of let-7 miRNAs, particularly let-7f-5p. They also demonstrate an apparent selectivity of SFN-mediated let-7f-5p and let-7g-5p induction towards the cancerous Caco-2 cell line, over the non-cancerous CCD-841 line. Such selectivity of let-7f-5p and/or let-7g-5p induction may contribute to SFN's potential to favour cytotoxicity towards cancerous cells over that towards healthy cells as reported in several contexts, and/or may serve to counterbalance the Nrf2-mediated cytoprotective response to moderate-dose SFN that inevitably occurs in tumour cells, whilst facilitating the cytoprotection of healthy cells.

Of additional interest is the observed SFN-mediated downregulation of miR-29b-3p in Caco-2 cells. The role of miR-29b-3p appears to be ambiguous, since both oncogenic (26) and tumour-suppressive (27) functions are reported. MiR-29b-3p is computationally predicted by miRanda to target the following tumour suppressor genes (TSGs): *ADAMTS9*, *ARRDC3*, *CASP7*, *FEM1B*, *HBPI*, *HPGD*, *ING3*, *NAV3*, *PTEN*, *RhoBTB1*, *SEL1L*, *SLC5A8*, *SUV420H2*, *TET1*, *TET2* and *UVRAG* (24). Direct interactions between miR-29b-3p and the 3'-UTRs of colorectal cancer-suppressing *TET1* and *TET2* have been demonstrated by luciferase assays already (28), (29), (30). Interestingly, a different ITC, phenethyl isothiocyanate, was reported to inhibit the upregulation of miR-29b-3p in mouse lung tissue mediated by environmental cigarette smoke (31). The ability of ITCs such as SFN and phenethyl isothiocyanate to modulate miR-29b-3p levels might be at least partially attributable to the induction of Nrf2, which itself has been reported to downregulate miR-

29b-3p, via the binding at and consequent repression of transcription from, the miR-29b-1 genetic locus (32). Therefore, experiments to test the ability of miR-29b-3p to bind to these predicted binding sites in the 3'-UTRs of these genes could lead to further insight into SFN's potential to upregulate tumour suppressor gene expression in colorectal cancer.

In conclusion, SFN upregulates let-7f-5p and let-7g-5p in Caco-2 cells but not in CCD-841. Let-7f-5p is able to bind directly to loci in the 3'-UTRs of CDC25A, HMGA2 and MYC that are computationally predicted as let-7f-5p binding sites, and may thus repress the translation of these oncogenes. The upregulation of let-7f-5p by SFN may thereby contribute to its suppressive effects on colorectal cancer.

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*. 2012;380(9859):2095-128.
2. World Health Organization: International Agency for Research on Cancer. Colorectal Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012 2016 [17 Oct 2016]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.
3. World Cancer Research Fund International. Colorectal cancer statistics 2015 [cited 2016 7 Sep 16]. Available from: <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/colorectal-cancer-statistics>.
4. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 Lyon, France: International Agency for Research on Cancer; 2010 [cited 2013 23rd August]. Available from: <http://globocan.iarc.fr>.
5. Smith D, Ballal M, Hodder R, Soin G, Selvachandran SN, Cade D. Symptomatic presentation of early colorectal cancer. *Annals of the Royal College of Surgeons of England*. 2006;88(2):185-90.
6. Tse G, Eslick GD. Cruciferous vegetables and risk of colorectal neoplasms: a systematic review and meta-analysis. *Nutr Cancer*. 2014;66(1):128-39.
7. Bacon JR, Williamson G, Garner RC, Lappin G, Langouet S, Bao Y. Sulforaphane and quercetin modulate PhIP-DNA adduct formation in human HepG2 cells and hepatocytes. *Carcinogenesis*. 2003;24(12):1903-11.
8. Cheung KL, Kong AN. Molecular targets of dietary phenethyl isothiocyanate and sulforaphane for cancer chemoprevention. *AAPS J*. 2010;12(1):87-97.
9. Zanichelli F, Capasso S, Bernardo G, Cipollaro M, Pagnotta E, Carteni M, et al. Low concentrations of isothiocyanates protect mesenchymal stem cells from oxidative injuries, while high concentrations exacerbate DNA damage. *Apoptosis*. 2012;17(9):964-74.

- 384 10. Higgins LG, Kelleher MO, Eggleston IM, Itoh K, Yamamoto M, Hayes JD.
385 Transcription factor Nrf2 mediates an adaptive response to sulforaphane that protects
386 fibroblasts in vitro against the cytotoxic effects of electrophiles, peroxides and redox-
387 cycling agents. *Toxicology and Applied Pharmacology*. 2009;237(3):267-80.
- 388 11. Sestili P, Paolillo M, Lenzi M, Colombo E, Vallorani L, Casadei L, et al.
389 Sulforaphane induces DNA single strand breaks in cultured human cells. *Mutat Res*.
390 2010;689(1-2):65-73.
- 391 12. Rajendran P, Kidane AI, Yu TW, Dashwood WM, Bisson WH, Lohr CV, et al.
392 HDAC turnover, CtIP acetylation and dysregulated DNA damage signaling in colon
393 cancer cells treated with sulforaphane and related dietary isothiocyanates. *Epigenetics*.
394 2013;8(6):612-23.
- 395 13. Slaby O, Sachlova M, Brezkova V, Hezova R, Kovarikova A, Bischofova S, et
396 al. Identification of MicroRNAs Regulated by Isothiocyanates and Association of
397 Polymorphisms Inside Their Target Sites with Risk of Sporadic Colorectal Cancer.
398 *Nutrition and cancer*. 2013;65(2):247-54.
- 399 14. Xu P, Billmeier M, Mohorianu I, Green D, Fraser WD, Dalmay T. An
400 improved protocol for small RNA library construction using High Definition adapters.
401 *Methods in Next Generation Sequencing*. 2015;2(1).
- 402 15. Sorefan K, Pais H, Hall AE, Kozomara A, Griffiths-Jones S, Moulton V, et al.
403 Reducing ligation bias of small RNAs in libraries for next generation sequencing.
404 *Silence*. 2012;3(1):1-11.
- 405 16. University of California Santa Cruz. Sequence and Annotation Downloads
406 2017 [cited 2017 13 Jan 17]. Available from:
407 <http://hgdownload.cse.ucsc.edu/downloads.html#human>.
- 408 17. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation
409 and deep-sequencing data. *Nucleic Acids Res*. 2011;39(Database issue):D152-7.
- 410 18. Prufer K, Stenzel U, Dannemann M, Green RE, Lachmann M, Kelso J.
411 PatMaN: rapid alignment of short sequences to large databases. *Bioinformatics*
412 (Oxford, England). 2008;24(13):1530-1.
- 413 19. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and
414 quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621-8.
- 415 20. Beckers M, Mohorianu I, Stocks M, Applegate C, Dalmay T, Moulton V.
416 Comprehensive processing of high throughput small RNA sequencing data including quality

- 417 checking, normalization and differential expression analysis using the UEA sRNA
418 Workbench. *RNA* 23 (6) pp. 823-835 (2017).
- 419 21. Li J, Tibshirani R. Finding consistent patterns: a nonparametric approach for
420 identifying differential expression in RNA-Seq data. *Statistical methods in medical*
421 *research*. 2013;22(5):519-36.
- 422 22. Mohorianu I, Schwach F, Jing R, Lopez-Gomollon S, Moxon S, Szittya G, et
423 al. Profiling of short RNAs during fleshy fruit development reveals stage-specific
424 sRNAome expression patterns. *Plant J*. 2011;67(2):232-46.
- 425 23. Barret T, Suzek T, Troup D, Wilhite S, Ngau W-C, Ledoux P, et al. NCBI
426 GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res*.
427 2005;33(1):D562-D6.
- 428 24. microRNA.org - Targets and Expression [Internet]. 2010. Available from:
429 <http://www.microrna.org>.
- 430 25. Iliopoulos D, Hirsch HA, Struhl K. An Epigenetic Switch Involving NF-κB,
431 Lin28, Let-7 MicroRNA, and IL6 Links Inflammation to Cell Transformation. *Cell*.
432 2009;139(4):693-706.
- 433 26. Adenuga D, Caito S, Yao H, Sundar IK, Hwang J-W, Chung S, et al. Nrf2
434 deficiency influences susceptibility to steroid resistance via HDAC2 reduction.
435 *Biochemical and Biophysical Research Communications*. 2010;403(3–4):452-6.
- 436 27. Wang B, Li W, Liu H, Yang L, Liao Q, Cui S, et al. miR-29b suppresses tumor
437 growth and metastasis in colorectal cancer via downregulating Tiam1 expression and
438 inhibiting epithelial-mesenchymal transition. *Cell Death Dis*. 2014;5:e1335.
- 439 28. Taylor MA, Wappett M, Delpuech O, Brown H, Chresta CM. Enhanced
440 MAPK signaling drives ETS1-mediated induction of miR-29b leading to
441 downregulation of TET1 and changes in epigenetic modifications in a subset of lung
442 SCC. *Oncogene*. 2016;35(33):4345-57.
- 443 29. Huang Y, Wang G, Liang Z, Yang Y, Cui L, Liu C-Y. Loss of nuclear
444 localization of TET2 in colorectal cancer. *Clinical Epigenetics*. 2016;8(1):9.
- 445 30. Cheng J, Guo S, Chen S, Mastriano Stephen J, Liu C, D'Alessio Ana C, et al.
446 An Extensive Network of TET2-Targeting MicroRNAs Regulates Malignant
447 Hematopoiesis. *Cell Reports*. 2013;5(2):471-81.

- 448 31. Izzotti A, Larghero P, Cartiglia C, Longobardi M, Pfeffer U, Steele VE, et al.
449 Modulation of microRNA expression by budesonide, phenethyl isothiocyanate and
450 cigarette smoke in mouse liver and lung. *Carcinogenesis*. 2010;31(5):894-901.

- 451 32. Shah NM, Zaitseva L, Bowles KM, MacEwan DJ, Rushworth SA. NRF2-
452 driven miR-125B1 and miR-29B1 transcriptional regulation controls a novel anti-
453 apoptotic miRNA regulatory network for AML survival. *Cell death and*
454 *differentiation*. 2015;22(4):654-64.

Figure 1: Normalized library abundances of let-7f-5p, let-7g-5p and miR-193b-3p in Caco-2 following 8 or 24 h SFN (10 μ M) treatment, or DMSO controls (0 h). Data are represented as the median abundance at each time point (of duplicates at 0 and 8 h, and of triplicates at 24 h), with the maximum and minimum values of replicate sets represented by the upper and lower bounds of error bars.

[A] Caco-2

[B] CCD-841

Figure 2: Northern Blot data for let-7f-5p and let-7g-5p in [A] Caco-2 and [B] CCD-841 cells, treated for 8 or 24 h with SFN (10 μ M) or DMSO controls. U6 was used as an internal control for normalization. Induction is expressed as [(miRNA band intensity)/(U6 band intensity)] for each sample relative to the control mean, and data are represented as means of duplicates $\pm$ (upper value – lower value)/2.

Figure 3: Northern Blot data for [A] miR-193b-3p and [B] miR-29b-3p in Caco-2 cells treated for 8 or 24 h with SFN (10 μ M) or DMSO controls. U6 was used as an internal control for normalization. Induction is expressed as $[(\text{miRNA band intensity})/(\text{U6 band intensity})]$ for each sample relative to the control mean, and data are represented as means of duplicates $\pm (\text{upper value} - \text{lower value})/2$.

Figure 4: Dot Blot data following the probing of a membrane that contained positive-sense oligonucleotides representing let-7a-5p, let-7b-5p, let-7c-5p, let-7d-5p, let-7e-5p, let-7f-5p, let-7g-5p and let-7i-5p, with probes complementary to either let-7f-5p, or to let-7g-5p.

Figure 5: Data from luciferase assays performed 48 h after co-transfecting HCT-116 cells with *pmiRGLO* vector and *let-7f-5p* mimic and/or *let-7f-5p* inhibitor and/or AllStars Negative Control siRNA. The *pmiRGLO* vector contained an insert consisting of four repeats of sequence antisense to that of *let-7f-5p*. Blank-corrected Firefly luminescence values were divided by blank-corrected Renilla luminescence values, and then ratios were normalized to controls (AllStars Negative Control siRNA at 200nM). Normalized data are represented as triplicate means $\pm$ SEM (* $p < 0.05$ according to two-tailed Student's T-test).

Figure 6: Data from luciferase assays performed 48 h after co-transfecting HCT-116 cells with *pmiRGLO* vector and *let-7f-5p* mimic and/or *let-7f-5p* inhibitor and/or AllStars Negative Control siRNA. The *pmiRGLO* vector contained an insertion of either part of the [A] *CDC25A*, [B] *HMGA2*, [C] *MYC* or [D] *KRAS* mRNA 3'-UTR containing (a) predicted *let-7f-5p* binding site(s), or a modified version of such in which said site(s) was/were mutated. Blank-corrected Firefly luminescence values were divided by blank-corrected Renilla luminescence values, and then ratios were normalized to controls (AllStars Negative Control siRNA at 200nM). Normalized data are represented as triplicate means $\pm$ SEM (* $p < 0.05$ according to two-tailed Student's *T*-test).