

The Roles of MicroRNAs in the Anti-Cancer Effects of Sulforaphane from Cruciferous Vegetables

A thesis submitted to the University of East Anglia
according to the requirements for the Degree of:

Doctor of Philosophy

By

Christopher A. Dacosta

Norwich Medical School



March 2017

This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Colorectal cancer is an increasingly important cause of mortality, whose incidence is inversely correlated with cruciferous vegetable consumption, from which can be obtained isothiocyanates such as sulforaphane. These are well-characterised regarding their cytoprotective and anti-tumour effects, thus believed to contribute to the observed diet-risk correlation; great interest lies in their potential use for chemoprevention and/or improvement of chemotherapy. While some mechanisms of action are well-established, such as their induction of antioxidant responses via the Nrf2 pathway, questions remain regarding its other mechanisms of action that appear to be vast and complex, more knowledge of which could shine a light upon means of safe and effective clinical application. There is evidence that sulforaphane can modulate the expression of microRNAs, which play major roles in development and disease by regulating gene expression, particularly in carcinogenesis. This study explored sulforaphane-induced microRNA modulation in adenocarcinoma Caco-2 and non-cancerous CCD-841 colorectal cells, and the potential role of such in the interactions between sulforaphane and colorectal carcinogenesis. The experimental system was initially validated by confirming the expected induction of Nrf2 and Nrf2-controlled genes at the mRNA and/or protein levels, by sub-cytotoxic doses of sulforaphane. Based upon data from the following miRNA-Seq-based expression profiling and individual microRNA assays, sulforaphane upregulated let-7f-5p and let-7g-5p expression in Caco-2 but not in CCD-841 cells, and upregulated miR-10a-5p and downregulated miR-193b-3p in Caco-2. The direct interaction of let-7f-5p with computationally-predicted mRNA-3'-UTR binding sites of *cell division cycle 25 homolog A* and *high-mobility group AT-hook 2* was confirmed by luciferase assays. Future experiments to confirm the effects of sulforaphane and/or let-7f-5p on these genes at the protein/mRNA levels could be informative, as could tests for synergistic and/or antagonistic interactions between sulforaphane and let-7f-5p. Insights from such experiments could eventually lead towards the development of more effective chemopreventative and chemotherapeutic strategies based upon sulforaphane and/or specific microRNA mimics/inhibitors.

Contents

Abstract	2
Contents	3
List of Abbreviations	12
List of Figures	19
List of Tables	42
Acknowledgements	45
Chapter 1: Introduction	46
1.1 Cruciferous Vegetables and Isothiocyanates.....	47
1.1.1 Cruciferous Vegetables	47
1.1.2 Isothiocyanates and Redox Status	47
1.2 Colorectal Cancer	50
1.2.1 Pathogenesis	52
1.2.1.1 Classic Vogelstein Model.....	52
1.2.1.2 Other Mechanisms.....	53
1.3 Isothiocyanates, Nuclear Factor (Erythroid-Derived 2)-Like 2 (Nrf2) and Metabolism	56
1.3.1 Nuclear Factor (Erythroid-Derived 2)-Like 2 (Nrf2).....	56
1.3.2 Phase I and II Enzymes	59
1.3.3 Mechanisms of Nrf2 Induction by Isothiocyanates	62
1.3.4 Isothiocyanates and Nrf2: Two Double-Edged Swords?	63
1.4 Epigenetics	68
1.5 MicroRNAs.....	70
1.5.1 Biogenesis.....	70
1.5.2 Canonical Model of Mature MicroRNA Activity	74
1.5.3 IsomiRs	76
1.5.4 Alternative Effects and Mechanisms of Activity	79

1.6	Links between ITCs, MicroRNAs, Biochemical Pathways and Cancer	80
1.6.1	Reported Modulation of MicroRNAs by Isothiocyanates.....	80
1.6.1.1	Examples of MicroRNA Modulation by ITC-Influenced Factors	83
1.6.2	Complexities in the Regulation of MicroRNA Abundance and Activity	88
1.6.2.1	Interfering RNAs.....	88
1.6.2.2	Interfering Proteins.....	89
1.6.2.3	Post-transcriptional MicroRNA Modifications	90
1.6.2.4	General Regulation	91
1.6.3	Interactions between MicroRNAs and Nrf2.....	92
1.6.4	MicroRNAs in Colorectal Cancer	92
1.6.4.1	MicroRNA Interactions with Vogelstein Model Pathogenesis	93
1.6.4.2	Other Potential Interactions between MicroRNAs and Colorectal Cancer.....	94
1.6.4.3	MicroRNA Interactions with the Warburg Effect	94
1.6.5	Methods to Detect Differences in MicroRNA Expression between Samples.....	95
1.6.5.1	MicroRNA Array Chip Hybridisation	95
1.6.5.2	MicroRNA Array RT-qPCR.....	95
1.6.5.3	MicroRNA Library Cloning and Deep Sequencing	95
1.6.5.4	Comparison of Methods	96
1.6.5.5	Challenges in MicroRNA Cloning: Adapter Ligation Bias ...	97
1.7	Aims and Objectives	98
Chapter 2: Methods.....		102
2.1	Cell Culture	102

2.2	Western Blotting to Assay Proteins.....	103
2.2.1	Cell Treatments to Assay the Effects of SFN on Proteins	103
2.2.2	Total Protein Isolation.....	103
2.2.3	Nuclear Protein Isolation	103
2.2.4	SDS-PAGE and Immunoblotting	104
2.2.4.1	Fluorescent Detection Protocol.....	104
2.2.4.2	Chemiluminescent Detection Protocol	105
2.3	MicroRNA Library Construction and Sequencing.....	106
2.3.1	Cell Treatments.....	106
2.3.2	Total RNA Extraction.....	106
2.3.3	MicroRNA Library Construction.....	106
2.3.3.1	Total RNA Purification	107
2.3.3.2	RNA Denaturation.....	107
2.3.3.3	3' HD Adapter Ligation.....	107
2.3.3.4	Degradation of Excess 3' HD Adapter	107
2.3.3.5	5' HD Adapter Ligation.....	108
2.3.3.6	cDNA Synthesis.....	108
2.3.3.7	PCR Amplification.....	108
2.3.3.8	PAGE Separation and Elution.....	109
2.3.4	Deep Sequencing.....	109
2.3.5	Data Analysis (Performed by Claudia Paicu)	110
2.4	Northern Blotting to Validate the Effects of SFN on MicroRNA Expression.....	111
2.4.1	Cell Treatments.....	111
2.4.2	Total RNA Extraction.....	111
2.4.3	Small RNA Extraction.....	112
2.4.4	RNA Electrophoresis and Blotting	112

2.4.5	Probe Hybridisation and Signal Detection	113
2.4.6	Dot Blots to Test Probe Cross-Reactivity.....	114
2.5	TaqMan® RT-qPCR to Assay the Effects of SFN on Messenger RNA	114
2.5.1	Cell Treatments	114
2.5.2	RNA Extraction	115
2.5.3	Reverse Transcription.....	115
2.5.4	Primers and Probes for TaqMan® RT-qPCR.....	115
2.5.5	RT-qPCR	116
2.5.6	Data Analysis.....	116
2.6	MicroRNA Assays by TaqMan® RT-qPCR	117
2.6.1	Cell Treatments	117
2.6.2	Small RNA Extraction	117
2.6.3	Total RNA Extraction	117
2.6.4	Reverse Transcription.....	117
2.7	Cloning 3'-UTRs into Vectors for Luciferase Assays.....	118
2.7.1	Making the Wild-Type 3' UTR Clones.....	118
2.7.2	Growing the Wild-Type Vectors in <i>E. coli</i>	121
2.7.3	Testing Prepared Vectors for Presence of 3'-UTR Inserts in Correct Orientation	122
2.7.4	Making Vectors with Mutated MicroRNA-Target Sites.....	123
2.8	Transfection and Luciferase Reporter Assays.....	124
2.8.1	Transfection.....	124
2.8.2	Luciferase Reporter Assays.....	125
2.9	Clonogenicity Assays	126
2.9.1	First Method.....	126
2.9.2	Second Method	126
2.10	Cell Viability Assays.....	127

2.11	Wound Healing Assays to Assess Cell Migration	128
2.12	Statistics	128
2.13	Rationale for Choice of Methods	129
2.13.1	RNA Extraction Methods.....	129
2.13.1.1	Organic Extraction	129
2.13.1.2	Filter-Based Extraction Methods.....	129
2.13.1.3	Direct Lysis	129
2.13.1.4	Choices of Methods.....	130
2.13.2	Protein Concentration Determination	130
2.13.3	Protein Expression Assay Methods	131
2.13.4	Measurement of Cell Viability.....	131
Chapter 3: Establishing and Validating the Experimental System		132
3.1	General Introduction	132
3.2	Effects of Sulforaphane and Hydrogen Peroxide on Cell Viability in Caco-2 and CCD-841	133
3.2.1	Sulforaphane.....	133
3.2.2	Hydrogen Peroxide.....	135
3.2.3	Sulforaphane vs. Hydrogen Peroxide.....	137
3.2.4	Methylselenocysteine (SeMSC)	137
3.3	Effects of Sulforaphane on Nrf2 and Nrf2-Controlled Phase II Enzyme Expression in Caco-2 and CCD-841 Cells.....	138
3.3.1	Thioredoxin Reductase 1	138
3.3.1.1	Messenger RNA Level.....	138
3.3.1.2	Protein Level With/Without Supplemental Selenium.....	140
3.3.2	UDP Glucuronosyltransferase 1 Family, Polypeptide A Cluster 147	
3.3.3	Nrf2 Nuclear Accumulation.....	149
3.4	Effects of Sulforaphane on Oncogenic Characteristics	151

3.4.1	Effects on Cell Migration.....	151
3.4.2	Effects on Clonogenicity	155
3.5	Conclusions.....	157
Chapter 4:	Profiling MicroRNA Expression	158
4.1	General Introduction.....	158
4.2	Caco-2 and CCD-841 MicroRNA Library Deep Sequencing	158
4.2.1	MicroRNAs Modulated by Sulforaphane Treatment in Caco-2 According to Libraries	161
4.2.2	MicroRNAs Modulated by Sulforaphane Treatment in CCD-841 According to Libraries	164
4.2.3	Examination of the Library Data Analysis and IsomiRs	166
4.3	Validating Differential MicroRNA Expression in Response to Sulforaphane.....	168
4.3.1	The Let-7 Family.....	169
4.3.1.1	Let-7f-5p in Caco-2	169
4.3.1.2	Let-7g-5p in Caco-2	170
4.3.1.3	Let-7f-5p and Let-7g-5p in CCD-841.....	172
4.3.1.4	Dot-blots Probed with Anti-Let-7f-5p Probe.....	176
4.3.1.5	Dot blots Probed with Anti-Let-7g-5p Probe.....	176
4.3.2	The MiR-10 Family	177
4.3.2.1	MiR-10a-5p in Caco-2.....	178
4.3.2.2	MiR-10b-5p in Caco-2.....	180
4.3.3	MiR-29b-3p.....	181
4.3.4	MiR-193b-3p.....	184
4.3.5	Other MiRNAs	185
4.3.5.1	Caco-2	185
4.3.5.2	CCD-841	188
4.4	Summary.....	192

Chapter 5: MicroRNA Target Validation	194
5.1 General Introduction	194
5.2 Let-7f-5p	194
5.2.1 Selection of Predicted Target Genes.....	195
5.2.2 Luciferase Assays	195
5.2.2.1 Control Experiments	196
5.2.2.2 BACH1	197
5.2.2.3 CDC25A(a)	198
5.2.2.4 HMGA2(b).....	199
5.2.2.5 MYC.....	200
5.3 Conclusions	202
Chapter 6: General Discussion	204
6.1 Recapitulation of Study Rationale and Objectives	204
6.2 General Summary of Findings	205
6.2.1 Validation of the System.....	205
6.2.1.1 Cell Viability Assays.....	205
6.2.1.2 The Effects of Sulforaphane on Nrf2 and Nrf2-Mediated Gene Expression.....	206
6.2.2 Effects of Sulforaphane on Oncogenic Characteristics	206
6.2.2.1 Cell Migration: Wound Healing Assays.....	206
6.2.2.2 Clonogenic Assays	207
6.2.3 Profiling MicroRNA Expression Following Sulforaphane Treatment.....	208
6.2.3.1 MicroRNA Libraries.....	208
6.2.3.2 MicroRNAs that were Individually Confirmed to be Modulated	208
6.2.3.3 MicroRNAs Unable to be Validated and Possible Explanations	210

6.2.4	Validating Predicted MicroRNA Targets and Functions.....	211
6.2.4.1	Predicted MicroRNA-Target Genes.....	211
6.2.4.2	Luciferase Assays to Confirm MicroRNA-3'-UTR Interactions 212	
6.3	MicroRNA Library Construction: Technical Challenges.....	214
6.3.1	Initial Set of Caco-2 Libraries.....	214
6.3.2	New Sets of Libraries for Caco-2 and CCD-841	215
6.4	MicroRNA Library Data vs. Findings in Literature	217
6.4.1	Inconclusive Observations.....	217
6.4.2	Undetected MicroRNAs	217
6.5	Future Perspectives	218
6.5.1	Possible Future Experiments.....	220
6.5.1.1	Effects of let-7f-5p Mimic and Inhibitor on Target Expression 220	
6.5.1.2	Other Potential Let-7 Targets and Sulforaphane-Modulated MicroRNAs	220
6.5.1.3	Crosslinking, Ligation and Sequencing of Hybrids.....	223
6.5.1.4	Xenograft Studies in Mice	224
6.5.1.5	Potential Clinical Applications of MicroRNA Manipulation	226
6.5.1.6	Clinical Applications of Sulforaphane	228
Appendix.....		229
7.1	Additions to Chapter 1.....	229
7.1.1	MicroRNA Nomenclature.....	229
7.1.1.1	Introduction of the -5p/-3p Suffix	229
7.1.1.2	IsomiRs	229
7.2	Additions to Chapter 4.....	231
7.2.1	MicroRNA Library Construction	231
7.2.1.1	Checking RNA Integrity.....	231

7.2.1.2	Elution of the Libraries	232
7.2.1.3	Quality Check and Quantification.....	242
7.2.2	MicroRNA Library Deep Sequencing Data Analysis.....	244
7.2.2.1	Size Class Distributions	244
7.2.2.2	Complexity Checks	249
7.2.3	Northern Blot/RT-qPCR Data	254
7.2.3.1	MicroRNAs Not Confirmed by Northern Blots/RT-qPCR .	254
7.3	Additions to Chapter 2 and Chapter 5.....	264
7.3.1	Construction of pmiRGLO Vectors with 3' UTRs.....	264
7.3.1.1	Cloning the 3' UTRs.....	264
7.3.1.2	Checking Prepared Plasmids for the Presence of Desired Inserts	271
7.3.1.3	Sequencing Data for the Wild-Type Vector Constructs....	275
7.3.1.4	Primer Design for PCR Mutagenesis of Predicted Target Loci	292
7.3.1.5	Checking Mutant Vectors for Successful Binding Site Mutations by Digestion.....	294
7.3.1.6	Sequencing Data for the Mutant Vectors	299
7.3.2	Luciferase Assay Optimisation	310
7.3.2.1	Firefly Data (Blank-Corrected)	310
7.3.2.2	Renilla Data (Blank-Corrected)	310
7.3.2.3	Firefly/Renilla	310
	Bibliography.....	313

List of Abbreviations

18q	chromosome 18 long arm
18q21	band 1 of region 2 of 18 q
3'-UTR	3'-untranslated region
5'-TOP	5'-terminal oligopyrimidine
5'-UTR	5'-untranslated region
AAD	5'-adapter-3'-adapter dimer
ACF	aberrant crypt focus
ADAR	adenosine deaminase
AGO	Argonaute
APC	adenomatous polyposis coli
ARE	antioxidant response element
ATP	adenosine triphosphate
BITC	benzyl isothiocyanate
CDC25A	cell division cycle 25 homolog A
cDNA	complementary DNA
CEBP α	CCAAT/enhancer-binding protein α
CEBP β	CCAAT/enhancer-binding protein β
ceRNA	competing endogenous RNA
ChIP-Seq	chromatin immunoprecipitation sequencing
CIN	chromosomal instability
circRNA	circular RNA

CLASH	crosslinking, ligation and sequencing of hybrids
CML	chronic myeloid leukaemia
CRC	colorectal cancer
Crm1	chromosomal maintenance 1
CYP	cytochrome P450
DCC	deleted in colorectal carcinoma
DGCR8	DiGeorge syndrome critical region 8
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNMT	DNA methyltransferase
EDC	1-ethyl-3-(dimethylaminopropyl) carbodiimide
EDVs	EDV™ Nanocells
ELISA	enzyme-linked immunosorbent assay
EMT	epithelial-mesenchymal transition
ER	endoplasmic reticulum
ERK	extracellular signal-regulated kinase
FAM	6-carboxyfluorescein
FBS	foetal bovine serum
FXR1	fragile X mental retardation-related protein 1
GCN	glucosinolate
GSH	glutathione
GSK3β	glycogen synthase kinase 3β

GSS	gene-specific sequence
GST	glutathione S-transferase
H ₂ O ₂	hydrogen peroxide
HDAC	histone deacetylase
HIF-1 α	hypoxia-inducible factor 1- α
HMGA2	high-mobility group AT-hook 2
HMT	histone methyltransferase
HNE	4-hydroxynonenal
HO-1	haem oxygenase 1
IC ₅₀	half maximal inhibitory concentration
IFN	interferon
IGF-1	insulin-like growth factor 1
IGF-1R	insulin-like growth factor-1 receptor
IL	interleukin
IRE1 α	inositol-requiring enzyme 1 α
IRES	internal ribosomal entry site
IRS-1	insulin receptor substrate
ITC	isothiocyanate
JAK	Janus kinase
Keap1	Kelch like-ECH-associated protein-1
KLF4	Kruppel-like factor 4
KSRP	KH-type splicing protein

LB	lysogeny broth
Ldbr	lariat debranching enzyme
LEDC	less economically developed country
lncRNA	long non-coding RNA
LOH	loss of heterozygosity
LPS	lipopolysaccharide
MAPK	mitogen-activated protein kinase
MEDC	more economically developed country
MGMT	O6-alkylguanine DNA alkyltransferase
miRNA	microRNA
MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2
MMR	mismatch repair
MPC	mitochondrial pyruvate carrier
mRNA	messenger RNA
MSI	microsatellite instability
mTOR	mammalian target of rapamycin
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NF-κB	nuclear factor κ-light-chain-enhancer of activated B cells
NQO1	NADPH:quinone reductase 1
Nrf2	nuclear factor (erythroid-derived 2)-like 2
NTA	non-templated addition
OD	optical density

PBS	phosphate-buffered saline
PBS-T (x%)	PBS supplemented with x% Tween 20 detergent
PCR	polymerase chain reaction
PDHX	pyruvate dehydrogenase protein X component
PEG	polyethylene glycol
PEITC	phenethyl isothiocyanate
PFA	paraformaldehyde
PGAM5	phosphoglycerate mutase family member 5 protein threonine phosphatase
PI3K	phosphoinositide 3-kinase
PI3K-AKT	phosphoinositide 3-kinase-RAC-alpha serine/threonine- protein kinase
PKC	protein kinase C
PLEKHA7	pleckstrin homology domain-containing family A member 7
PMF1	polyamine-modulation factor 1
PNPT1	polyribonucleotide nucleotidyltransferase 1
PPAR γ	proliferator-activated receptor γ
PRE	polyamine-responsive element
pre- miRNA	precursor miRNA
pri- miRNA	primary miRNA transcript

PTEN	protein phosphatase and tensin homologue deleted on chromosome 10
PTENP1	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)
PVDF	polyvinylidene difluoride
qPCR	quantitative PCR
RISC	RNA-induced silencing complex
RNase	endoribonuclease
ROS	reactive oxygen species
R-Smad	receptor-regulated mothers against decapentaplegic
RT-qPCR	reverse transcription-quantitative polymerase chain reaction
SCID	severe combined immunodeficient
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SeMSC	methylselenocysteine
SFN	sulforaphane
SFRP	secreted frizzled-related protein
SG	stress granule
SIRT1	sirtuin 1
SMAD2	mothers against decapentaplegic homolog 2
SMAD4	mothers against decapentaplegic homolog 4
sMAF	small MAF
SSAT	spermidine/spermine N1-acetyltransferase

STAT	signal transducer and activator of transcription 3
TAMRA	tetramethylrhodamine
TBE	tris/borate/EDTA
TGF β	transforming growth factor β
TM	melting temperature
TNF- α	tumour necrosis factor- α
TrxR-1	thioredoxin reductase 1
TSG	tumour suppressor gene
UGT	uridine 5'-diphospho-glucuronosyltransferase
VSS	vector-specific sequence
YAP	Yes-associated protein
ZEB	zinc-finger enhancer binding

List of Figures

Figure 1: The general structural formulae of the glucosinolates and isothiocyanates are shown. The hydrolysis of a glucosinolate by myrosinases can produce an isothiocyanate. Different members of the glucosinolates and isothiocyanates differ at the positions indicated by the letter 'R' 48

Figure 2: A summary of colorectal cancer pathogenesis according to the Vogelstein model. 55

Figure 3: A diagram illustrating the interaction of ROS and electrophiles with the Nrf2-Keap1 complex, and the subsequent induction of Nrf2 activity and reducing (antioxidant) enzyme expression. (1) ROS and electrophiles interact with cysteine residues of Keap1, inducing modifications that weaken Keap1's association with Nrf2. (2) Nrf2 thus avoids proteasomal degradation, and can translocate to the nucleus. (3) Nrf2 binds to the *ARE* which is found in the 5' flanking region of many genes, some of which act to shift cellular redox balance in a reductive direction..... 59

Figure 4: A diagram illustrating the metabolism of xenobiotics by phase I and phase II enzymes, and the influence of the ITC SFN upon the activities of each. SFN can inhibit CYP enzymes, and also induce phase II enzymes such as GST- α , HO-1, NQO-1 and TrxR-1. This shifts the balance between phase I and II metabolism to inhibit the accumulation of the undesirable intermediate metabolites..... 61

Figure 5: A diagram illustrating the process of ITC-GS conjugation in the cell cytoplasm, and the release of ITCs in the plasma. (1) ITCs are conjugated to GSH under the action of GST enzymes. (2) ITC-GS conjugates are exported from the cell. (3) In the plasma, ITC-GS conjugates are hydrolysed, resulting in the regeneration of ITC molecules. (4) The ITC molecules can enter the same cell, or another (Zhang 2000)..... 62

Figure 6: The acute oxidative effect induced by a hormetic dose of ITCs contributes to their ability to induce Nrf2, and ultimately generates a net medium-term reducing effect in cells. 63

Figure 7: Small regions of the genome that are relatively rich in CpG dinucleotides are called CpG islands. CpG islands have a strong tendency to be found in a methylated state unless located proximal to gene promoters. 69

Figure 8: A diagram illustrating the recruitment of HDACs to methylated cytosine residues, and the subsequent repression of gene transcription. (1) Methylated cytosine residues are bound by the methyl-CpG-binding domain of an HDAC, recruiting the HDAC to the site. (2) The HDAC removes acetyl groups, exposing positive charge on the histone tails, which interact with the negatively charged phosphate backbone of DNA. This causes the local chromatin to become more compact, thus sterically inhibiting the binding of transcription factors and machinery.69

Figure 9: A diagram illustrating the canonical pathway of miRNA expression in animal cells. (1) A genetic locus is transcribed by RNA polymerase II, producing a pri-miRNA, which is several hundred nucleotides long. (2) The pri-miRNA is bound by DGCR8, which recruits Drosha, which cleaves the pri-miRNA into pre-miRNAs that are about 70 nucleotides long and have 2-nucleotide overhangs at their 3' ends. (3) The 3' 2-nucleotide overhangs are recognised by Exportin 5, which uses GTP to transport them from the nucleus to the cytoplasm. (4) Dicer recognises the same 3' overhang and makes a nick in the loop region of the pre-miRNA, generating an imperfectly paired linear RNA duplex, each strand of which bears a 3' 2-nucleotide overhang. (5) The linear RNA duplex is unwound; one strand remains associated with Dicer as the mature miRNA, and becomes part of a RISC upon association with AGOs.72

Figure 10: A diagram illustrating the generation of a pre-miRNA from a lariat structure, itself excised from the intronic region of an early mRNA transcript. (1) A genetic sequence is transcribed by RNA polymerase II. (2) The intronic sequence of the RNA transcript is excised in the form of a lariat structure. (3) Ldbp1 converts the lariat structure into a pre-miRNA..... 74

Figure 11: A diagram illustrating the interaction between a mature miRNA, an AGO protein, and a target mRNA. The AGO interacts with the miRNA and at the 5' and 3' ends, via its PIWI and PAZ domains respectively. A 5' seed

region of the miRNA base pairs with part of the 3' untranslated region of a target mRNA, to which it is partially or fully complementary. Subsequently, the AGO can interact with the target mRNA to inhibit its translation. 76

Figure 12: An illustration of how the generation of isomiRs may help to intensify the desirable regulation of a specific target, without the undesirable intensification of regulation of another target. (A) The amplification of the canonical miRNA increases the regulation of the desired mRNA target, but also undesirably intensifies the regulation of an 'off-target' mRNA. (B) The generated isomiRs all contribute to the desirable regulation of the target mRNA, but do not necessarily share the same 'off-target' mRNAs, thus potentially undesirable effects on 'off-target' mRNAs are diluted..... 78

Figure 13: A graph illustrating the increase in the number of mature human miRNAs listed in miRBase with each new version, according to the miRBase Blog..... 97

Figure 14: The effects of SFN on Caco-2 cell viability at 24 h and 48 h. Caco-2 cells were treated for 24 or 48 h, by adding to culture medium in sextuplicate, DMSO-diluted SFN at various concentrations, or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Cell viability was measured by the MTT assay and evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (*P<0.05)..... 133

Figure 15: The effect of SFN on CCD-841 cell viability at 24 and 48 h. CCD-841 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, DMSO-diluted SFN at various concentrations, or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Cell viability was measured by the MTT assay and evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (*P<0.05)..... 134

Figure 16: The effects of H₂O₂ on Caco-2 cell viability at 24 h and 48 h. Caco-2 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, water-dissolved H₂O₂ at various concentrations, or water

(control). Cell viability was evaluated as ($OD_{670} - OD_{570}$) at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$)...... 135

Figure 17: The effect of H_2O_2 on CCD-841 cell viability at 24 h and 48 h. CCD-841 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, water-dissolved H_2O_2 at various concentrations, or water (control). Cell viability was evaluated as ($OD_{670} - OD_{570}$) at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$)...... 136

Figure 18: The effects of SeMSC on Caco-2 cell viability at 24 h. Caco-2 cells were treated for 24 h by adding to culture medium in sextuplicate, water-dissolved SeMSC at various concentrations, or water (control). Cell viability was evaluated as ($OD_{670} - OD_{570}$) at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$)...... 138

Figure 19: The effects of 24 h SFN treatment on the TrxR-1 mRNA transcript level in Caco-2 cells. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (2.5, 5, 10 or 20 μ M) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Relative TrxR-1 mRNA levels were assayed via TaqMan RT-qPCR as described in Methods, using 18S as an internal control. Relative TrxR-1 expression is evaluated by the Pfaffl method from C_T values, taking into account efficiencies derived from standard curve data. Data are represented as mean \pm S.E.M. of triplicates (* $P < 0.05$)...... 139

Figure 20: The effects of 24 h SFN treatment on the TrxR-1 mRNA transcript level in CCD-841 cells. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (2.5, 5, 10 or 20 μ M) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Relative TrxR-1 mRNA levels were assayed via TaqMan RT-qPCR as described in Methods, using 18S as an internal control. Relative TrxR-1 expression is evaluated by the Pfaffl method from C_T values, taking into account efficiencies derived from standard curve data. Data are represented as mean \pm S.E.M. of triplicates (* $P < 0.05$)...... 140

Figure 21: The effect of 48 h SFN treatment on TrxR-1 protein levels in Caco-2 cells. Cells were treated for 48 h by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2. ... 142

Figure 22: The effect of SFN treatment on TrxR-1 protein levels in Caco-2 cells, with or without selenium supplementation. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10µM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM). Relative protein abundance was assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2. ... 143

Figure 23: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated for 24 h by adding to culture medium, DMSO-diluted SFN (2.5 or 5µM) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are representative of a single experiment. 144

Figure 24: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated for 48 h

by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2. 145

Figure 25: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10µM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM) to culture medium. Relative protein abundance was assayed via Western Blotting as described in Methods, using β-actin as an internal control. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2. 146

Figure 26: The effect of 24 h SFN treatment on UGT1A protein levels in Caco-2 cells. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Relative protein abundance was assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2. 147

Figure 27: The effect of 72 h SFN treatment on UGT1A protein levels in Caco-2 cells. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control) (final DMSO

concentrations were 0.05% (v/v)). Relative protein abundance was assayed via Western Blotting as described in Methods, using β -actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as $[(\text{TrxR-1 band intensity})/(\text{Internal control band intensity})]$ for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2..... 148

Figure 28: The effect of SFN treatment on nuclear accumulation of Nrf2 in Caco-2 cells. Cells were treated for 4, 8, 12 or 24 h with SFN (10 μ M) or for 4 h with SFN (5 or 20 μ M) by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Relative protein levels were assayed via Western Blotting as described in Methods, using Sam68 as an internal control for normalisation. Induction is evaluated and illustrated in the graph as $[(\text{Nrf2 band intensity})/(\text{Internal control band intensity})]$ for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2..... 150

Figure 29: Caco-2 cells were seeded at densities sufficient that confluent monolayers were formed by 24 h. Scratch wounds were created in confluent monolayers using a pipette tip, and cells were then treated by adding to culture medium in triplicate SFN (5 to 20 μ M) or DMSO; final DMSO concentrations were 0.05% (v/v). Cells were subsequently incubated for 72 h, then fixed, stained and photographed (with 5 photographs along the vertical axis of each scratch wound). The 4th image of every wound is shown. The average gap width between leading cells was calculated in each image, then data across all 5 images per well were averaged. Data from image analyses are represented in the graph as the means of three biological replicates \pm S.E.M..... 152

Figure 30: CCD-841 cells were seeded so as to form confluent monolayers at 24 h. Scratch wounds were then created in confluent monolayers with a pipette tip, and cells were then treated by adding to culture medium in triplicate SFN (5 to 20 μ M) or DMSO; final DMSO concentrations were 0.05% (v/v). Cells were subsequently incubated for 48 h, then fixed, stained and

photographed (with 5 photographs along the vertical axis of each scratch wound). The 4th image of every wound is shown. The average gap width between leading cells was calculated in each image, then data across all 5 images per well were averaged. Data from image analyses are represented in the graph as the means of three biological replicates \pm S.E.M. (*P<0.05).
 154

Figure 31: Caco-2 cells were seeded at low density (400 cells/well) in 6-well plates (surface area/well = 9.5cm²) and then incubated for 48 h prior to treatment. After 48 h, cells were treated by adding to culture medium in sextuplicate, DMSO-diluted SFN (5-20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Cells were then incubated for a further 14 days, after which cells were fixed and stained, and the number of colonies in each well was counted under an inverted microscope. Absolute colony counts are presented in the table and relative colony formation is illustrated in the graph, where data are represented as means of six biological replicates \pm S.E.M. (*P<0.05). 156

Figure 32: Caco-2 cells were cultured to 70-80% confluence in 24-well plates (surface area/well = 1.9cm²), then treated with DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO; final DMSO concentrations were 0.05% (v/v). Subsequently, cells were detached using trypsin and then re-seeded at low density (400 cells/well) in 6-well plates (surface area/well = 9.5cm²). Cells were then incubated for a further 14 days after which cells were fixed and stained, and the number of colonies in each well was counted. Relative colony formation is illustrated in the graph, where data are represented as means of six biological replicates \pm S.E.M. 157

Figure 33: Normalised read numbers for hsa-let-7f-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 169

Figure 34: The effect of SFN treatment on the level of hsa-let-7f-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO

concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 170

Figure 35: Normalised read numbers for hsa-let-7g-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 171

Figure 36: The effect of SFN treatment on the level of hsa-let-7g-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 172

Figure 37: Normalised read numbers for hsa-let-7f-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 173

Figure 38: Normalised read numbers for hsa-let-7g-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 173

Figure 39: The effect of SFN treatment on the level of hsa-let-7f-5p in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed

via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 174

Figure 40: The effect of SFN treatment on the level of hsa-let-7g-5p in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 175

Figure 41: A test of the cross-reactivity of the probe anti-sense to hsa-let-7f-5p with other let-7 family members. Positive sense oligomers composed of the sequences of each miRNA were blotted onto a membrane, which was then probed against hsa-let-7f-5p. 176

Figure 42: A test of the cross-reactivity of the probe anti-sense to hsa-let-7g-5p with other let-7 family members. Positive sense oligomers composed of the sequences of each miRNA were blotted onto a membrane, which was then probed against hsa-let-7g-5p. 176

Figure 43: Normalised read numbers for hsa-miR-10a-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 177

Figure 44: The effect of SFN treatment on the level of hsa-miR-10a-5p in Caco-2 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via TaqMan RT-qPCR as described in Methods,

using U6 as an internal control for normalisation. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (* $p < 0.05$)..... 178

Figure 45: The effect of 24 h SFN treatment on the level of hsa-miR-10a-5p in Caco-2 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control for normalisation. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (* $p < 0.05$). 179

Figure 46: The effect of SFN treatment on the level of hsa-miR-10a-5p[[isomiR]]21_42 in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 180

Figure 47: Normalised read numbers for hsa-miR-10b-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 181

Figure 48: The effect of SFN treatment on the level of hsa-miR-29b-3p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 182

Figure 49: Normalised read numbers for hsa-miR-29b-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 183

Figure 50: Normalised read numbers for hsa-miR-193b-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 184

Figure 51: The effect of SFN treatment on the level of hsa-miR-193b-3p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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. 185

Figure 52: Normalised read numbers for hsa-miR-182-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 186

Figure 53: Normalised read numbers for hsa-miR-192-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 187

Figure 54: Normalised read numbers for hsa-miR-17-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.... 187

Figure 55: Normalised read numbers for hsa-miR-1296-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 188

Figure 56: Normalised read numbers for hsa-miR-182-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 189

Figure 57: Normalised read numbers for hsa-miR-106b-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 190

Figure 58: Normalised read numbers for hsa-miR-181a-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 191

Figure 59: Normalised read numbers for hsa-miR-345-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars. ... 192

Figure 60: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insert consisting of four repeats of sequence antisense to that of let-7f-5p. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as

triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).
..... 197

Figure 61: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the BACH1 3'-UTR containing two predicted let-7f-5p binding sites, or a modified form of this 3'-UTR portion with the predicted binding sites mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test). 198

Figure 62: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the CDC25A 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test). 199

Figure 63: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-

7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the HMGA2 3'-UTR containing three predicted let-7f-5p binding sites, or a modified form of this 3'-UTR portion with the predicted binding sites mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (*p<0.05 according to two-tailed Student's T-test)..... 200

Figure 64: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the MYC 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (*p<0.05 according to two-tailed Student's T-test)..... 201

Figure 65: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the KRAS 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase

Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).....202

Figure 66: A sketch illustrating the hypothesis that the let-7f-5p site-binding dynamics are distinctly non-linear and differ between different 3'-UTRs. In the hypothetical situation here illustrated, the basal effective let-7f-5p concentration means that the CDC25A and MYC binding sites are already saturated with let-7f-5p but the BACH1 and HMGA2 sites are largely unbound. Thus, addition of let-7f-5p mimic has little to no effect on CDC25A and MYC, but strong effects on BACH1 and HMGA2, whilst the addition of let-7f-5p inhibitor has little to no effect on BACH1 and HMGA2 but significantly decreases the occupancy of CDC25A and MYC sites by let-7f-5p.213

Figure 67: Images of SYBR Gold-stained PCR PAGE gels for Caco-2 libraries; one constructed by the old protocol (A), and the other by the new protocol (B). Although prominent AAD bands are apparent in both gels, their intensities relative to those of the higher miRNA bands are much lower in the gel image for the library constructed by the new protocol than in that for that constructed by the old protocol.216

Figure 68: A diagram illustrating potentially oncogenic miRNAs that are predicted by miRanda (with mirSVR scores ≤ -1.2) to be targeted by let-7f-5p and/or let-7g-5p.221

Figure 69: A diagram illustrating potentially tumour-suppressive miRNAs that are predicted by miRanda (with mirSVR scores ≤ -1.2) to be targeted by miR-29b-3p or miR-193b-3p.223

Figure 70: Ethidium bromide-stained images of the agarose gel through which Caco-2 total RNA samples were electrophoresed, in order to check

RNA integrity prior to library construction. The samples subsequently used for library construction are annotated below the gel image.....	231
Figure 71: Ethidium bromide-stained image of the agarose gel through which CCD-841 total RNA samples were electrophoresed, in order to check RNA integrity prior to library construction. The samples subsequently used for library construction are annotated below the gel image.	232
Figure 72: Images of ethidium bromide-stained gels with PAGE-separated PCR products of cDNA that was reverse transcribed from small RNAs captured from samples by the 5' and 3' HD adapters. The bands visible between 140 and 160bp are the desired microRNA bands, which were cut out and from which libraries were eluted, while those visible below are from the adapter-adapter dimers, which care was taken to avoid.	242
Figure 73: Ethidium bromide-stained gel with Caco-2 libraries electrophoresed for quality checking and quantification of concentrations.	242
Figure 74: The first of two ethidium bromide-stained gels with CCD-841 libraries electrophoresed for quality checking and quantification of concentrations.....	243
Figure 75: The second of two ethidium bromide-stained gels with CCD-841 libraries electrophoresed for quality checking and quantification of concentrations.....	244
Figure 76: A graph representing the Caco-2 size class distribution data for A1, A2 and A3.....	245
Figure 77: A graph representing the Caco-2 size class distribution data for B1, B2 and B3.....	246
Figure 78: A graph representing the Caco-2 size class distribution data for C1, C2 and C3.	246
Figure 79: A graph representing the CCD-841 size class distribution data for X1, X2 and X3.....	247

Figure 80: A graph representing the CCD-841 size class distribution data for Y1, Y2 and Y3.	248
Figure 81: A graph representing the CCD-841 size class distribution data for Z1, Z2 and Z3.	248
Figure 82: A graph illustrating the complexity distribution data for A1, A2 and A3.	249
Figure 83: A graph illustrating the complexity distribution data for B1, B2 and B3.	250
Figure 84: A graph illustrating the complexity distribution data for C1, C2 and C3.	250
Figure 85: A graph illustrating the complexity distribution data for X1, X2 and X3.	252
Figure 86: A graph illustrating the complexity distribution data for Y1, Y2 and Y3.	252
Figure 87: A graph illustrating the complexity distribution data for Z1, Z2 and Z3.	253
Figure 88: The effect of SFN treatment on the level of hsa-miR-10b-5p[[isomiR]]26_47 in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10µM) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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 ± (upper value – lower value)/2.	254
Figure 89: The effect of SFN treatment on the level of hsa-miR-10b-5p in Caco-2 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (10µM) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative	

miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (*p<0.05)..... 255

Figure 90: The effect of 24 h SFN treatment on the level of hsa-miR-10b-5p in Caco-2 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (*p<0.05). 255

Figure 91: The effect of SFN treatment on the level of hsa-miR-17-3p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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. 256

Figure 92: The effect of SFN treatment on the level of hsa-miR-182-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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. 257

Figure 93: The effect of SFN treatment on the level of hsa-miR-192-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in

duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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)/2.258

Figure 94: The effect of SFN treatment on the level of hsa-miR-1296-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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)/2.259

Figure 95: The effect of SFN treatment on the level of hsa-miR-106b-5p in CCD-841 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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)/2.260

Figure 96: The effect of 24 h SFN treatment on the level of hsa-miR-345-5p in CCD-841 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as

described in Methods, and data are represented as means of triplicates \pm S.E.M. (*P<0.05)..... 261

Figure 97: The effect of SFN treatment on the level of hsa-miR-181a-5p in CCD-841 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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. 262

Figure 98: A test of the cross-reactivity of the probe anti-sense to hsa-miR-181a-5p with hsa-miR-181b-5p. Positive sense oligomers representing the sequence of each miRNA were blotted onto a membrane, which was then probed against hsa-miR-181a-5p..... 263

Figure 99: The effect of SFN treatment on the level of hsa-miR-182-5p in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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. 264

Figure 100: A diagram illustrating the expected cleavage of the pmiRGLO vector by the HindIII restriction enzyme. 271

Figure 101: An ethidium bromide-based image of a 1% agarose gel into which HindIII digestion products for BACH1, BACH2, BACH3, BACH4, CDC25A[b] 1, CDC25A[b] 2, CDC25A[b] 3, CDC25A[b] 4, KRAS 1, KRAS 2, KRAS 3 and KRAS 4 were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2 h. 272

Figure 102: An ethidium bromide-based image of a 1% agarose gel into which HindIII digestion products for BACH5, BACH6, BACH7, BACH8, BACH9 and BACH10 were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2h.273

Figure 103: An ethidium bromide-based image of a 1% agarose gel into which the HindIII digestion products of the HMGA2[a] 1, HMGA2[a] 2, HMGA2[a] 3, HMGA2[a] 4, HMGA2[b] 1, HMGA2[b] 2, HMGA2[b] 3, HMGA2[b] 4, MYC 1, MYC 2, MYC 3 and MYC 4 plasmids were loaded along a 1kb ladder (NEB), and electrophoresed at 100V for 2 h 30 min. ...274

Figure 104: An ethidium bromide-based image of a 1% agarose gel into which the HindIII and NheI digestion products of the CDC25A[a] 5, CDC25A[a] 6, CDC25A[a] 7, CDC25A[a] 8, CDC25A[a] 9, MYC 1, MYC 2, MYC 3 and MYC 4 plasmids were loaded alongside 1kb and 100bp ladders (NEB), and electrophoresed at 80V for 1 h.275

Figure 105: A diagram (not to scale) illustrating the loci of the existing restriction sites in the BACH1 construct, and those of new restriction sites created by successful seed region mutagenesis294

Figure 106: A diagram (not to scale) illustrating the loci of the existing restriction sites in the CDC25Aa construct, and those of new restriction sites created by successful seed region mutagenesis294

Figure 107: A diagram (not to scale) illustrating the loci of the existing restriction sites in the CDC25Ab construct, and those of new restriction sites created by successful seed region mutagenesis295

Figure 108: A diagram (not to scale) illustrating the loci of the existing restriction sites in the HMGA2a construct, and those of new restriction sites created by successful seed region mutagenesis295

Figure 109: A diagram (not to scale) illustrating the loci of the existing restriction sites in the HMGA2b construct, and those of new restriction sites created by successful seed region mutagenesis296

Figure 110: A diagram (not to scale) illustrating the loci of the existing restriction sites in the KRAS construct, and those of new restriction sites created by successful seed region mutagenesis..... 296

Figure 111: A diagram (not to scale) illustrating the loci of the existing restriction sites in the MYC construct, and those of new restriction sites created by successful seed region mutagenesis..... 297

Figure 112: An ethidium bromide-based image of a 1% (w/v) agarose gel into which the EcoRV digestion products of BACH1 and HMGA2b first-site mutants, and the BamHI digestion products of CDC25Ab, HMGA2a and KRAS first-site mutants were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2h. 298

Figure 113: An ethidium bromide-based image of a 1% (w/v) agarose gel into which the BamHI digestion products of BACH1 and HMGA2b double mutants, and the MfeI digestion product of CDC25Ab double mutant were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 60V for 1h..... 299

Figure 114: The luminescent signals from the experiments to optimise the amounts of DNA and Attractene transfection reagent to use are represented in this graph as means \pm S.E.M. 311

Figure 115: The [Firefly luminescence]/[*Renilla* luminescence] ratios are represented as means \pm S.E.M..... 312

List of Tables

Table 1: A summary of miRNAs reported to be altered by 48 h SFN treatment in NCM356 or NCM460 cells according to microRNA low-density array RT-qPCR experiments (Slaby, Sachlova et al. 2013), along with potential functions reported in the literature.	81
Table 2: A table illustrating some of the interactions between HDACs/DNMTs, inflammation, Nrf2 and redox status.....	86
Table 3: The sequences of the primers and probes used for RT-qPCR.....	115
Table 4: The sequences of the designed PCR primers.....	119
Table 5: A summary of where and what mutations were directed to.	123
Table 6: The identity of each Caco-2 miRNA library, the proportion of reads of each remaining after the quality check and removal of adapter-adapter reads, and the subsequent proportion of reads mapping to the human genome. Libraries were constructed with three biological replicates per time point; 'A' libraries represent controls, 'B' 8 h SFN treatment, and 'C' 24 h SFN treatment.....	159
Table 7: The identity of each CCD-841 miRNA library, the proportion of reads of each remaining after the quality check and removal of adapter-adapter reads, and the subsequent proportion of reads mapping to the human genome. Libraries highlighted in red were excluded from further analysis because very low proportions of their reads were mappable to the human genome, potentially due to contamination of samples with non-human RNA during the construction of these libraries. Libraries were constructed with three biological replicates per time point; 'X' libraries represent controls, 'Y' 8 h SFN treatment, and 'Z' 24 h SFN treatment.	160
Table 8: A summary of the numbers of total and differentially expressed known miRNAs detected in each cell line.....	161

Table 9: The identities and normalised read number medians for miRNAs in Caco-2 for which the calculated fold change of expression between 0 and 8 h, 8 and 24 h, or 0 and 24 h of SFN (10µM) treatment was ≥ 1.9 . $vA = \text{Median}((\text{Reads in A1}),(\text{Reads in A2}))$; $vB = \text{Median}((\text{Reads in B1}),(\text{Reads in B2}))$; $vC = \text{Median}((\text{Reads in C1}),(\text{Reads in C2}),(\text{Reads in C3}))$	161
Table 10: The identities and normalised read number medians for miRNAs in CCD-841 for which the calculated fold change of expression between 0 and 8 h, 8 and 24 h, or 0 and 24 h of SFN (10µM) treatment was ≥ 1.9 . $vX = \text{Median}((\text{Reads in X1}),(\text{Reads in X2}),(\text{Reads in X3}))$; $vY = (\text{Reads in Y3})$; $vZ = \text{Median}((\text{Reads in Z1}),(\text{Reads in Z2}))$	164
Table 11: The canonical sequences of hsa-let-7a-5p, hsa-let-7f-5p and hsa-let-7g-5p.....	168
Table 12: A summary of the size class distributions of reads from the sequencing of the Caco-2 libraries.....	245
Table 13: A summary of the size class distributions of reads from the sequencing of the CCD-841 libraries.	247
Table 14: A summary of the complexity scores at each sequence length, in each Caco-2 library.....	249
Table 15: A summary of the complexity scores at each sequence length, in each CCD-841 library.	251
Table 16: A summary of the expected HindIII digestion fragments.	271
Table 17: A summary of the expected HindIII digestion products. *The HMGA2[b] region has one HindIII digestion site within it.	273
Table 18: A summary of expected products following digestion with both HindIII and NheI.	274
Table 19: A list of the full sequences of forward and reverse primers designed for the PCR mutagenesis of predicted miRNA-targeted loci in 3'-UTRs.....	292

Acknowledgements

I am extremely grateful to my primary and secondary supervisors, Dr. Yongping Bao and Professor Tamas Dalmay, for their expert support and guidance throughout the research and completion of the thesis.

I would like to thank the Faculty of Medicine and Health Sciences at UEA for providing the studentship that enabled me to undertake my PhD studies at the University.

I am indebted to Claudia Paicu and Irina-Ioana Mohorianu from the School of Computing Sciences for the quality-checking, annotation and analysis of the microRNA library deep sequencing data.

I would like to warmly thank all members of the Bao and Dalmay lab groups, including Wei Wang, Ping Xu and Guy Wheeler for their extensive technical guidance and support, particularly regarding the construction of the microRNA libraries, and the cloning of human mRNA 3'-UTRs into vectors for luciferase assays. I would also like to thank all of my lab and office colleagues that have contributed to providing a warm and friendly working environment over the years.

Being able to continue my musical hobbies by being a member of the UEA Music Society has provided me with a much-needed counterbalance to work life throughout the course of my studies, thus I would like to thank all members and committee of the UEA Music Society (2013 – 2017). I am also deeply indebted to the clergy and congregation of St. Thomas Church, Norwich, without whose friendship and support I would have been unable to complete this thesis.

Lastly, but not least, I am deeply grateful to my parents and grandparents for their unconditional love, support, reassurance and prayers.

Chapter 1: Introduction

Links between an individual's dietary and lifestyle habits and their risk of developing cancers have long been observed, and are highlighted by various epidemiological studies (Huxley, Ansary-Moghaddam et al. 2009). In particular, inverse correlations between the intake of fruits and vegetables and the risk of developing chronic diseases such as cancers have been suggested since the 1950s. Various explanations have been sought for this apparent connection, ranging from the dietary displacement of cancer-promoting foods and improved micronutrient intake, to (more recently) the direct activity of non-nutritive yet bioactive compounds found in fruits and vegetables, referred to as phytochemicals. Phytochemicals are believed to play a number of roles in plants, from imparting colour and flavour, to acting as endogenous pesticides and fungicides. The potential benefits to human health of such phytochemicals have become increasingly apparent over recent decades.

Globally, colorectal cancer (CRC) was the third most commonly diagnosed cancer in 2012 according to the World Cancer Research Fund International, accounting for 10% of total cancer diagnoses (World Cancer Research Fund International 2015). The incidence of CRC by country is positively associated with economic development, and is most notably on the rise in countries experiencing rapid economic development and/or increased adoption of 'Western' dietary habits. As is also the case for other cancers, epidemiological studies link various lifestyle and dietary factors to one's risk of developing CRC (Hagggar and Boushey 2009).

The isothiocyanates (ITCs) are a particular group of phytochemicals that are obtained by the consumption of cruciferous vegetables, and to which the relatively strong inverse correlations between cruciferous vegetable intake and cancer risk have often been attributed. Interest in the potential of ITCs to directly inhibit the development of chronic diseases has intensified since the 1980s.

1.1 Cruciferous Vegetables and Isothiocyanates

1.1.1 Cruciferous Vegetables

As mentioned above, epidemiological studies have shown an inverse correlation between the consumption of cruciferous vegetables - such as broccoli, cabbage and Brussel sprouts - and the risk of developing various cancers, such as those of the lung (Lam, Gallicchio et al. 2009) and colon (Tse and Eslick 2014), independently of overall vegetable intake. Found particularly in the cruciferous variety of vegetables are phytochemicals called glucosinolates (GCNs), which are broken down upon vegetable consumption to form several types of product, including ITCs. ITCs *per se* have been extensively studied *in vitro* for their potential to inhibit the proliferation, migration and invasiveness of tumour cell lines, whilst promoting cell cycle arrest and apoptosis. Additionally, they have been shown to have chemopreventative potential in animal carcinogenesis models, thought to be largely due to the preservation of macromolecular integrity. Together, these studies imply that ITCs are highly likely to be one, if not the only, causative factor for the observed inverse correlation between cruciferous vegetable intake and cancer risk.

1.1.2 Isothiocyanates and Redox Status

ITCs are believed to have evolved to defend the plants against microbial and small animal (mostly insect) attack. The breakdown of the biologically inactive GCNs is mediated by the action of hydrolytic myrosinase enzymes. Intact plant tissues separately compartmentalise the GCNs in S-cells (Koroleva, Davies et al. 2000) and myrosinases in myrosin cells (Andréasson, Jørgensen et al. 2001), thus preventing the myrosinase-mediated hydrolysis of the GCNs. Only when the cells are ruptured e.g. by food processing and/or chewing, or by microbial infection, are the myrosinases able to interact with, and thus hydrolyse, the GCNs. This hydrolysis has the potential to form different types of compound, including ITCs, thiocyanates and nitriles (Burow, Bergner et al. 2007). Of these, the ITCs are to date the best characterised for their bioactivity.

The GCNs and ITCs are each groups of related compounds; a specific ITC is typically generated by the hydrolysis of a specific GCN. Broccoli and broccoli sprouts are abundant in a particular GCN called glucoraphanin, whose hydrolysis gives rise to the ITC sulforaphane (SFN). SFN *per se* has been frequently studied for its potential to inhibit the development of various chronic health conditions such as osteoarthritis (Davidson, Jupp et al. 2016), cardiovascular diseases (Bai, Wang et al. 2015), autism (Singh, Connors et al. 2014) and cancers (Clarke, Dashwood et al. 2008), which is believed to be largely attributable to its broad anti-inflammatory, antioxidant and chemoprotective properties.

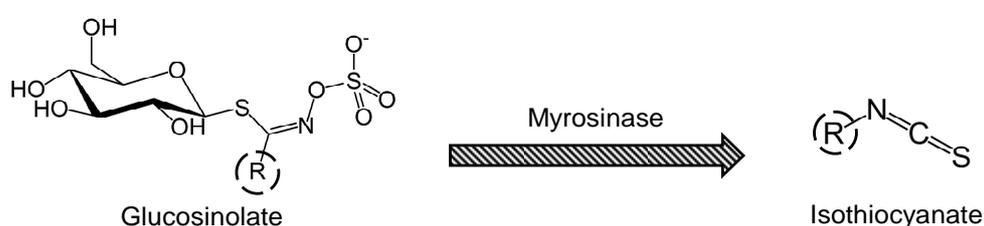


Figure 1: The general structural formulae of the glucosinolates and isothiocyanates are shown. The hydrolysis of a glucosinolate by myrosinases can produce an isothiocyanate. Different members of the glucosinolates and isothiocyanates differ at the positions indicated by the letter 'R'.

The mechanisms underlying the multitude of reported anti-cancer effects of ITCs *in vitro* and *in vivo* are likely to be wide-ranging, complex and interlinked. One of the longest-established effects of ITCs, as of some other phytochemicals, is the induction of nuclear factor (erythroid-derived 2)-like 2 (Nrf2): a transcription factor that activates many genes coding for reductive (antioxidant) proteins that reduce oxidative reactive oxygen species (ROS), thereby inhibiting oxidative stress i.e. a deleteriously excessive shift of cellular redox status towards oxidation (Li and Kong 2009).

Reductive i.e. antioxidant effects have traditionally been thought of as overwhelmingly beneficial to human health, due to the potential of ROS to react non-specifically with and thereby damage macromolecules such as carbohydrates, lipids, proteins and nucleic acids. Such reactions promote inflammatory signalling that is linked to the pathogenesis of chronic

diseases such as atherosclerosis, neurodegenerative disorders and cancers, as well as the general age-related decline of organ function. ROS can also disrupt homeostatic cell signalling, promote the apoptotic death of neurons and hepatocytes, and increase the frequency of potentially carcinogenic mutations, due to DNA oxidation. In fact, a hypothesis popular in the middle of the 20th century pinned ROS as the direct cause of the 'ageing process', and proposed that ageing might be halted by artificially eliminating ROS (Harman 1956).

However, it has since come to light that ROS are not simply undesirable by-products of aerobic metabolism as previously characterised. Specific ROS such as the superoxide anion and hydrogen peroxide play vital roles in cell signalling and immune cell activity, while oxidative shifts in redox status are important for the promotion of cell growth and differentiation in particular contexts. An example of the importance of ROS in cell signalling is the dependence of pancreatic β -cell insulin secretion upon the glucose uptake-induced generation of ROS, as demonstrated by the fact that a deficiency of ROS impairs glucose-stimulated insulin secretion (Pi, Zhang et al. 2010). James Watson accordingly hypothesised in 2014 that a deficiency of ROS actually contributes to the development of type 2 diabetes. On the other hand, excessive oxidation inevitably promotes pancreatic β -cell death, so balance is crucial. Another example of ROS-mediated signalling is the promotion of protein kinase C (PKC) activity induced by the binding of hydrogen peroxide to PKC regulatory domains (Gopalakrishna and Gundimeda 2002). PKC activation leads to the upregulation of anti-apoptotic and proliferation-stimulating signalling pathways.

Although ITCs are often described as having reducing (antioxidant) effects, they are not directly reductive – their acute effects in the cells are actually oxidative because they deplete intracellular reductants such as glutathione (GSH). In the medium term, however, cellular responses to ITCs are such that their net effects tend to be 'U-shaped', in that they lead to a reduction of ROS at moderate doses, while significantly elevating them at very high doses. This makes ITCs particularly interesting with regards to their potential

for protecting non-cancerous cells whilst simultaneously exerting toxicity towards cancer cells. This will be discussed in further detail below.

1.2 Colorectal Cancer

CRC – also known as colon, bowel or rectal cancer - originates from the oncogenic transformation of epithelial cells that line the colon or rectum. The rising global incidence of CRC was about 1.4 million in 2012 (World Health Organization: International Agency for Research on Cancer 2016), and in 2010, 715 000 mortalities were directly attributable to CRC: 46% more than were so in 1990 (Lozano, Naghavi et al. 2012). CRC was the third most commonly diagnosed cancer in 2012 according to the World Cancer Research Fund International, as mentioned above.

Reported CRC incidence is considerably higher in more economically developed countries (MEDCs) than in less economically developed countries (LEDCs) (Ferlay, Shin et al. 2010), and is rising most notably in countries experiencing rapid economic growth and/or cultural adoption of ‘Westernised’ dietary habits. One expects the incidence of chronic diseases such as cancers to naturally increase as mortality from communicable diseases drops, as tends to occur with economic development due to improvements in medical and public health infrastructure. Additionally, an increase in reported incidence does not directly imply an increase in actual incidence, since the former could be a result of improved diagnosis. Nevertheless, there is compelling evidence that a number of dietary and lifestyle factors do affect one’s risk of developing CRC, particularly concerning patterns of fruit, vegetable and meat consumption, fibre intake, and energy balance.

Typical warning signs of CRC are considered to include worsening constipation, reduced stool calibre, blood in stools, weight loss, loss of appetite, nausea/vomiting and/or anaemia, present in an individual aged 50 or over. Formal diagnosis is made through a colonoscopy or sigmoidoscopy of the bowel. The degree to which a cancer has spread is assessed via X-ray computed tomography scans of the chest, abdomen and pelvis, after which the tumour is graded according to the TNM Classification of Malignant

Tumours system based on spread, lymph node involvement, and the number of metastases.

Although early-stage colorectal tumours are now treatable with a high success rate, obvious symptoms of CRC tend not to present until the often-terminal late adenoma/malignant stage (Smith, Ballal et al. 2006). Thus, the screening of asymptomatic individuals is essential for the diagnosis of early stage CRCs. Under the NHS of the UK, individuals judged to be at high risk (based on age, family history and current conditions) are invited for routine screening, via faecal occult blood tests and sigmoidoscopies or colonoscopies. This screening strategy can detect CRC up to 3 years prior to the hypothetical emergence of tangible symptoms (Cunningham, Atkin et al. 2010). A screening programme for all individuals aged ≥ 55 years is also in the process of being implemented across the UK.

In addition to seeking early detection of CRC, there is increasing focus on prevention via dietary and lifestyle intervention strategies. According to Bowel Cancer UK, one's risk of developing CRC is reduced by:

- Eating a diet high in fibre, fruits and vegetables.
- Limiting red meat and alcohol consumption.
- Keeping physically active.

In recent years the focus on the potential use of phytochemicals from fruits, vegetables and herbs has intensified, both for chemoprevention in healthy individuals, and/or chemotherapeutic application against tumours in cancer patients. As mentioned previously, the consumption of fruits and vegetables has long been inversely associated with the development of chronic diseases such as cancers, neurodegenerative disorders and cardiovascular disease, often presumed to be owing to their content of micronutrients, phytochemicals, fibre, and/or dietary displacement of cancer-promoting foods. However, vegetables of the cruciferous variety are more strongly linked to reduced risks of cancer than vegetables in general. Cruciferous vegetables are distinct in that they contain GCNs, from which are derived ITCs, whose chemopreventative and/or chemotherapeutic potential has

become of particular interest over the past 30 years. The colorectal tissues are inevitably better predisposed than many others of the body to encounter orally ingested compounds prior to their hepatic metabolism (Pappa, Bartsch et al. 2007), thus observable effects of the compounds *in vitro* are more likely to bear physiological relevance.

1.2.1 Pathogenesis

1.2.1.1 Classic Vogelstein Model

Many models for the pathogenesis of CRC have been proposed, the most well-known of which is that proposed by Vogelstein (Fearon and Vogelstein 1990), which is summarised in Figure 2 and is as follows:

1. Adenomatous polyposis coli (APC) – a promoter of β -catenin degradation - becomes mutationally inactivated. β -catenin thus accumulates in the cytoplasm, then translocates to the nucleus where it transcriptionally upregulates several proto-oncogenes such as *MYC* that promote stem cell renewal and proliferation, leading to the development of an aberrant crypt focus (ACF).
2. The ACF continues to grow and its intracellular and intercellular organisation becomes more abnormal, eventually giving rise to an early adenoma. Constitutively-activating mutations occur in members of the *RAS* growth factor gene family, often *K-RAS* (Frank 2007), further promoting proliferation.
3. Several genes located in the long arm of chromosome 18 (18q) become deleted - particularly those in band 1 of region 2 of 18q (18q21), including important tumour suppressor genes (TSGs) such as *deleted in colorectal carcinoma (DCC)*, *Mothers against decapentaplegic homolog 2 (SMAD2)* and *Mothers against decapentaplegic homolog 4 (SMAD4)*. *DCC* promotes apoptosis in the absence of netrin-1 (Mazelin, Bernet et al. 2004). *SMAD2* and *SMAD4* are transforming growth factor beta (TGF β)-activated genes that may mediate TGF β 's tumour-suppressive effects in the context of CRC (Frank 2007). The loss of these TSGs facilitates the development of a late adenoma. TGF β is multifunctional and its net

role may become oncogenic once the pathways of TGF β -mediated growth inhibition, such as those mediated by *SMAD2* and *SMAD4*, are inactivated.

4. The loss of *TP53* is the final trigger for progression of the adenoma to malignant cancer (Frank 2007). Its protein, p53, is a crucial tumour suppressor that usually promotes apoptosis in response to DNA damage. Interestingly, it has also been shown to inhibit glycolysis and promote oxidative phosphorylation by increasing fructose-2,6-bisphosphate levels, thus opposing the proliferation-facilitating glycolytic metabolic switch observed in cancer cells (i.e. the Warburg effect) (Gonzalez, Alvarez et al. 2014).

Interestingly, SFN has been shown to inhibit β -catenin signalling and to suppress the growth of human breast cancer cell lines and xenograft models (Li, Zhang et al. 2010), thus indicating SFN's potential to inhibit the formation of ACFs and/or the development of early adenomas.

Chromosomal instability (CIN) results from the disruption of genes governing chromosome duplication and/or segregation. CIN can promote the loss of entire chromosomes and loss of heterozygosity (LOH) (Frank 2007), thus promoting the loss of important TSGs such as *APC*, therefore potentially triggering and/or accelerating classic Vogelstein pathogenesis.

1.2.1.2 Other Mechanisms

Microsatellite instability (MSI) is a general term used to describe the situation in which the pathways of DNA mismatch repair (MMR) are disrupted such that repetitive sequences - such as microsatellite regions - become significantly more prone to mutation, resulting in gene mutational inactivation and/or hyperactivation. MSI has been demonstrated to trigger hyperactivating mutations in the β -catenin gene itself, and thus initiate an *APC*-independent pathway of tumorigenesis. Interestingly, most CRCs exhibit either CIN or MSI, but not both (Frank 2007).

The initiators and facilitators of pathogenesis discussed so far are all genetic in nature. However, there is increasing evidence for the importance of

epigenetic changes in CRC pathogenesis, which could complement the genetic changes described in the Vogelstein model:

- The loss of APC activity facilitates enhanced β -catenin accumulation in the first step of pathogenesis according to the Vogelstein model. This is usually attributed to inactivating mutation, but hypermethylation of *APC* is also often observed in CRC vs. non-cancerous colorectal tissues (Hiltunen, Alhonen et al. 1997).
- *DCC*, a pro-apoptotic gene, is lost via the deletion of chromosomal region 18q21 in the transition from early to late adenoma according to the Vogelstein model. However, this pro-apoptotic gene also has a tendency to be hypermethylated in CRC cell lines (Derks, Bosch et al. 2009).
- *MutL homolog 1, colon cancer, nonpolyposis type 2 (MLH1)* is a gene that is important for effective MMR and the hypermethylation of its promoter can promote MSI, thus increasing the occurrence of potentially carcinogenic mutations.
- *O6-alkylguanine DNA alkyltransferase (MGMT)* is a gene that is important for the effective removal of mutagenic guanine adducts, and its hypermethylation has been demonstrated to play an important role in the promotion of mutation (Frank 2007).
- The secreted frizzled-related proteins (SFRPs) are a group of tumour-suppressive proteins that inhibit the accumulation of β -catenin, and their genes are frequently found to be hypermethylated in CRC samples (Lao and Grady 2011).

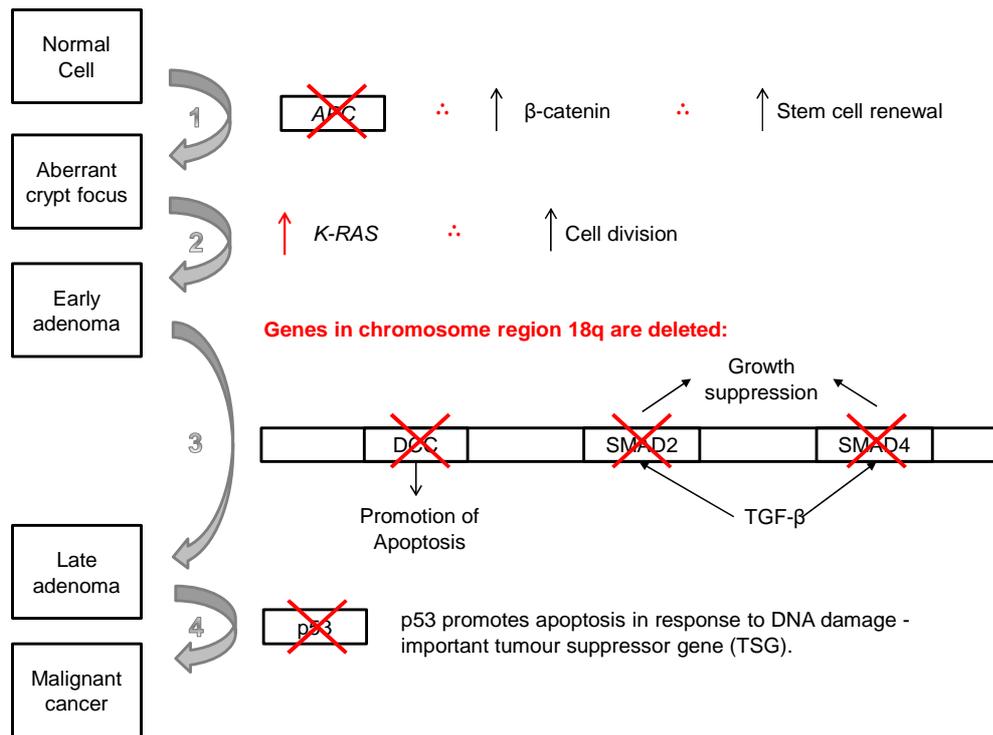


Figure 2: A summary of colorectal cancer pathogenesis according to the Vogelstein model.

Recent genomic studies have revealed that there are on average nine oncogenic mutations in protein-coding genes observable per CRC (Yamagishi, Kuroda et al. 2016). Such studies have not elucidated a panel of specific frequent driver mutations, but rather a panel of specific signalling pathways upon which various driver mutations tend to converge regarding their impact, namely: the Wnt, TGF β , phosphoinositide 3-kinase (PI3K), receptor tyrosine kinase-RAS and p53 signalling pathways (Yamagishi, Kuroda et al. 2016).

1.2.1.2.1 The Warburg Effect

As first observed and reported by Otto Heinrich Warburg in the early 20th century, the metabolism of tumour cells shifts such that a greater proportion of ATP is generated from glycolysis than from oxidative phosphorylation, despite an abundant supply of oxygen. It involves the metabolism of pyruvate being shifted to favour the generation of lactate rather than of acetyl-coenzyme A, which can be achieved by the suppression of mitochondrial

pyruvate entry, the inhibition of mitochondrial pyruvate metabolism to generate acetyl-coenzyme A, and/or the upregulation of enzymes that metabolise pyruvate to lactate. Efficient mitochondrial pyruvate entry is facilitated by the mitochondrial pyruvate carrier (MPC) proteins. The Warburg effect drastically reduces ATP generation per glucose, yet can actually increase ATP production per time given sufficient glucose abundance. Therefore, glucose uptake by tumour cells is characteristically high – a phenomenon that is exploited in the clinic to visualise tumours by positron emission tomography with fluorescently-labelled glucose analogues. Another consequence of the Warburg effect is the increased production of intermediates necessary for anabolism, which facilitates more rapid proliferation. Therefore, it is unsurprising that the Warburg effect can also be observed in rapidly-dividing healthy cells.

Interestingly, the genes coding for MPCs are often inactivated in CRC cell lines, in which their forced re-expression is able to inhibit anchorage-independent growth and suppress stem cell markers (Schell, Olson et al. 2014). Also interesting is that the expression of MPC1 has been shown to correlate positively with that of APC (Schell, Olson et al. 2014), and negatively with that of MYC (Schell, Olson et al. 2014). One may therefore postulate that interactions between dietary components and MPCs or other proteins influencing the balance between ATP-generating pathways might contribute to the observed links between diet and carcinogenesis.

1.3 Isothiocyanates, Nuclear Factor (Erythroid-Derived 2)-Like 2 (Nrf2) and Metabolism

1.3.1 Nuclear Factor (Erythroid-Derived 2)-Like 2 (Nrf2)

Nrf2 is a transcription factor best characterised for its role in defending cells against oxidative stress by transcriptionally upregulating reducing (antioxidant) enzymes. Its functions are now known to be vast in number and complexity, to extend far beyond redox homeostasis, and to include the modulation of cell proliferation, differentiation, cell fate determination and

stem cell maintenance in context- and cell-type-specific manners (Murakami and Motohashi 2015).

Basally, most of the Nrf2 in a given cell is located in cytoplasm where it is bound to a repressor protein called Kelch like-ECH-associated protein-1 (Keap1), whose binding both prevents the translocation of Nrf2 to the nucleus, and targets Nrf2 for 26S proteasomal degradation by recruiting the ubiquitin E3 ligase (Cullinan, Gordan et al. 2004). Keap1-Nrf2 binding is facilitated by phosphoglycerate mutase family member 5 protein threonine phosphatase (PGAM5), via which they are both tethered to the outer mitochondrial membrane (Xue, Momiji et al. 2015).

Oxidative stress triggers the release of Nrf2 from Keap1-mediated repression, partly via the oxidative modification of Keap1 at its Cys273 and Cys288 residues. Released Nrf2 is subsequently phosphorylated by casein kinase-2, which facilitates its translocation to the nucleus via importins $\alpha 5$ and $\beta 1$. Nuclear Nrf2 heterodimerises with small Maf (sMAF) proteins, with which it binds to the *antioxidant response element (ARE)* present in the 5' flanking regions of various Nrf2-activated genes, in a manner that tends to promote their transcription. Such genes include those coding for reducing (antioxidant) enzymes such as:

- Glutathione S-transferases (GSTs)
- Thioredoxin reductase 1 (TrxR-1)
- NADPH:quinone reductase 1 (NQO1)
- Haem oxygenase 1 (HO-1)
- Uridine 5'-diphospho-glucuronosyltransferases (UGTs)

Nuclear Nrf2 is eventually inactivated via acetylation. The phosphorylation of Nrf2 by Fyn kinase promotes its export from the nucleus by chromosomal maintenance 1 (Crm1). Fyn kinase itself can be activated upon interaction with the PGAM5 that is liberated and thus able to translocate to the nucleus when the cytoplasmic Nrf2-Keap1-PGAM5 complex is disrupted, thus forming part of a negative feedback loop (Xue, Momiji et al. 2015). The return of Nrf2 to the cytoplasm makes it once again liable to Keap1-directed

ubiquitination and degradation. Bach1 can inhibit the Nrf2-mediated transcriptional activation of *ARE*-regulated genes by competing with Nrf2 for sMAF binding; while Bach1-sMAF complexes are often able to bind to *AREs*, they tend to either not activate transcription or be less activating than Nrf2-sMAF complexes. However, Bach1 does not repress all Nrf2-targeted target genes equally; it has been demonstrated that Bach1 can significantly repress the Nrf2-induced expression of HO-1, but not that of TrxR-1 (Reichard, Motz et al. 2007). Bach1 can be directly inactivated by oxidation, and such is sufficient for the oxidative stress-mediated induction of HO-1 independently of cytoplasmic Nrf2 activation and nuclear translocation, because the basal nuclear levels of Nrf2 – albeit low - are sufficient to facilitate this.

As well as promoting the nuclear translocation of cytoplasmic Nrf2 protein, ROS can upregulate Nrf2 at both the transcriptional and translational levels. The Nrf2 messenger RNA (mRNA) transcript is one of many stress-responsive transcripts containing an internal ribosomal entry site (IRES) within the 5'-untranslated region (5'-UTR); the IRES enables translation to be initiated independently of cap-binding proteins (Li, Thakor et al. 2010). Oxidative stress can promote the translation of such cap-independent stress-responsive mRNAs, whilst simultaneously repressing the cap-dependent translation of typical mRNA transcripts (Li, Thakor et al. 2010).

The reductive (antioxidative) shift in redox status mediated by Nrf2 is most associated with cytoprotection, by helping to protect against oxidative damage to macromolecules. However, this redox shift also plays roles in cell fate determination. For example, an Nrf2-mediated reduction of ROS in macrophages can inhibit their differentiation to osteoclasts (Murakami and Motohashi 2015). Nrf2 can also influence cell fate determination in a redox-independent fashion, such as by transcriptionally upregulating CCAAT/enhancer-binding protein β (CEBP β) and proliferator-activated receptor γ (PPAR γ): two promoters of adipocyte differentiation (Murakami and Motohashi 2015). Nrf2 is continuously synthesised and degraded, so an increase in cellular Nrf2 levels often results from a synthesis-favouring perturbation of the basal steady-state synthesis-degradation balance.

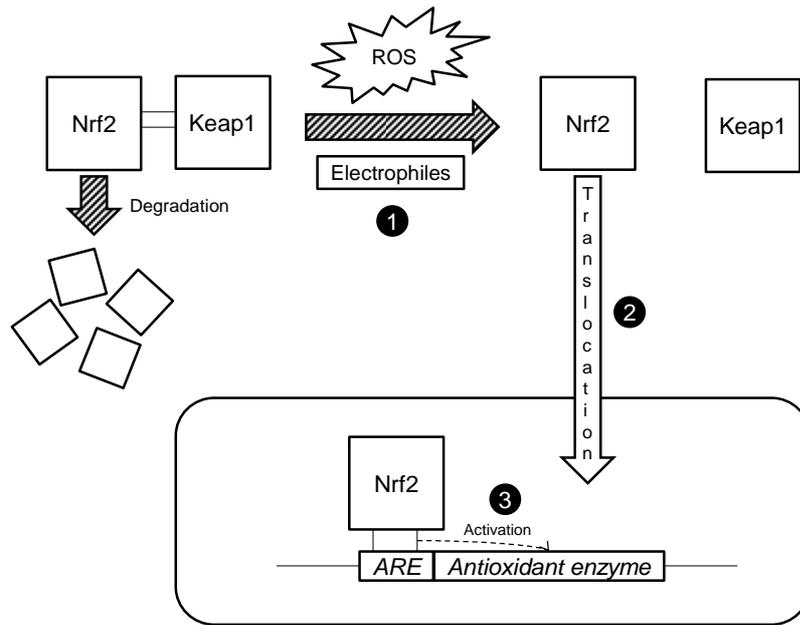


Figure 3: A diagram illustrating the interaction of ROS and electrophiles with the Nrf2-Keap1 complex, and the subsequent induction of Nrf2 activity and reducing (antioxidant) enzyme expression. (1) ROS and electrophiles interact with cysteine residues of Keap1, inducing modifications that weaken Keap1's association with Nrf2. (2) Nrf2 thus avoids proteasomal degradation, and can translocate to the nucleus. (3) Nrf2 binds to the ARE which is found in the 5' flanking region of many genes, some of which act to shift cellular redox balance in a reductive direction.

1.3.2 Phase I and II Enzymes

The breakdown of xenobiotics and endogenous hormones typically consists of several steps, the first being their metabolism by phase I enzymes, which include the cytochrome P450 (CYP) family of haemoproteins. These enzymes covalently modify their substrates by hydroxylation; their products tend to be highly reactive (often electrophilic or nucleophilic), and thus able to induce oxidative stress and macromolecular damage (Lewis 2002). The enzymes that metabolise these reactive intermediates are called phase II enzymes, and include the Nrf2-activated reducing enzymes described previously. The products of phase II enzyme-catalysed reactions tend to be hydrophilic, and thus liable to excretion.

An appropriate balance between phase I and phase II enzyme activity is crucial for limiting the accumulation of the reactive intermediates, and

consequent oxidative stress and macromolecular damage. In fact, there are environmental carcinogens that are not particularly detrimental *per se*, but whose intermediate metabolic products (i.e. the products of phase I metabolism) exhibit the associated deleterious, oncogenic effects.

Interestingly, ITCs have been demonstrated to inhibit the expression and/or activity of the CYP phase I enzyme family in certain contexts. For example, treatment of primary rat hepatocytes with aliphatic ITCs (40 μ M) was shown to repress both CYP1A1 and CYP1A2 at the transcriptional level (La Marca, Beffy et al. 2012). The combination of phase I enzyme-repression and phase II enzyme-induction (e.g. via Nrf2) by ITCs acts to stem the accumulation of cytotoxic and/or carcinogenic phase I metabolites.

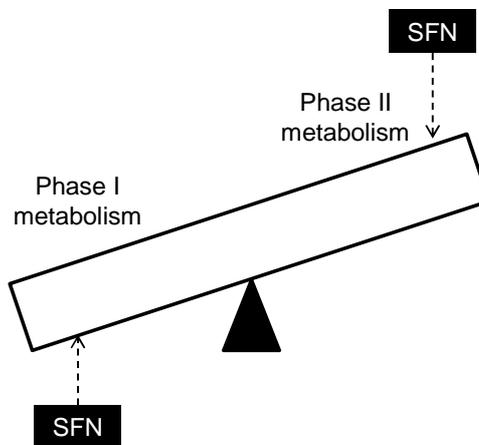
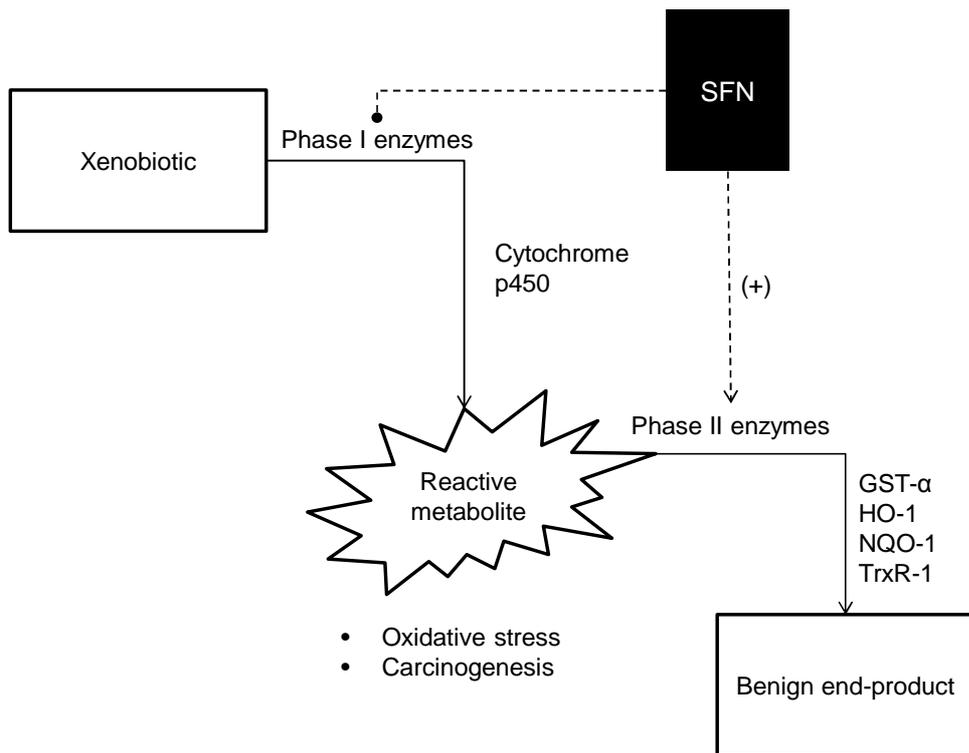


Figure 4: A diagram illustrating the metabolism of xenobiotics by phase I and phase II enzymes, and the influence of the ITC SFN upon the activities of each. SFN can inhibit CYP enzymes, and also induce phase II enzymes such as GST- α , HO-1, NQO-1 and TrxR-1. This shifts the balance between phase I and II metabolism to inhibit the accumulation of the undesirable intermediate metabolites.

1.3.3 Mechanisms of Nrf2 Induction by Isothiocyanates

Multiple and synergistically-acting mechanisms through which ITCs may induce Nrf2 have been elucidated, but there is no clear consensus on the most significant contributors. Although ITCs are sometimes referred to as antioxidants (reductants), their acute effects are actually oxidative, primarily due to depletion of the major intracellular electron donor protein GSH. GSH is conjugated to the ITCs by GSTs to form complexes that are exported from the cell (Higgins, Kelleher et al. 2009). Interestingly, SFN can also directly induce ROS generation in mitochondria, by inhibiting complex III of the mitochondrial respiratory chain, leading to the accumulation of ubisemiquinone, which donates electrons to molecular oxygen to generate the superoxide anion and hydrogen peroxide (Sestili, Paolillo et al. 2010).

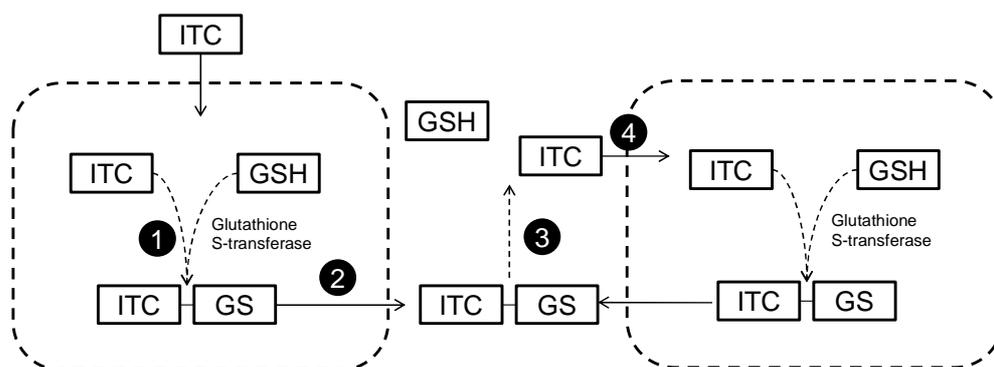


Figure 5: A diagram illustrating the process of ITC-GS conjugation in the cell cytoplasm, and the release of ITCs in the plasma. (1) ITCs are conjugated to GSH under the action of GST enzymes. (2) ITC-GS conjugates are exported from the cell. (3) In the plasma, ITC-GS conjugates are hydrolysed, resulting in the regeneration of ITC molecules. (4) The ITC molecules can enter the same cell, or another (Zhang 2000).

Inevitably, the acute oxidant effects of ITCs themselves contribute to the ability of ITCs to induce Nrf2 (Higgins, Kelleher et al. 2009). One hypothesis is that acute oxidative stress is the primary mediator of Nrf2 induction by ITCs, and that the medium-term antioxidant (reducing) effects observed following treatment with ITCs at moderate doses occur simply by the activity of Nrf2-induced reducing enzymes and increased GSH synthesis eventually outweighing the initial oxidant effects. In support of this, treatment of hepatocellular carcinoma HepG2-C8 cells with 25 μ M SFN was shown to

reduce cellular GSH levels at 4 h, but to increase them 2.2-fold at 24 h versus controls (Kim, Hu et al. 2003).

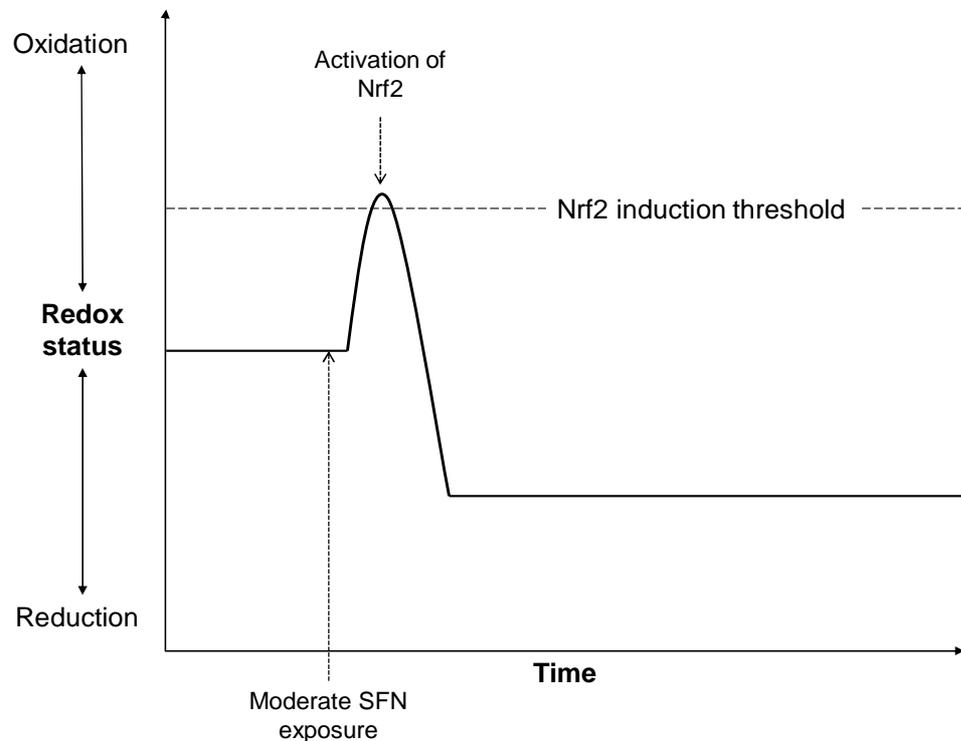


Figure 6: The acute oxidative effect induced by a hormetic dose of ITCs contributes to their ability to induce Nrf2, and ultimately generates a net medium-term reducing effect in cells.

On the other hand, it has been demonstrated that redox-independent effects of ITCs can contribute to Nrf2 activation, such as the inhibition of p38 mitogen-activated protein kinase (MAPK) activity. P38 MAPK can repress the activity of Nrf2 by phosphorylating it in a manner that promotes its binding to Keap1 (Keum, Yu et al. 2006).

1.3.4 Isothiocyanates and Nrf2: Two Double-Edged Swords?

The net medium-term effect of ITC exposure can be either oxidative or reductive (antioxidative), depending upon the dose to which a cell is exposed. At low-to-moderate doses of ITCs, the reductive effects of Nrf2 induction are observed to outweigh the acute oxidative effects, producing an overall medium-term reductive (antioxidative) effect. For this reason, ITCs are sometimes referred to as 'indirect antioxidants' and are frequently

observed to have 'antioxidant' effects. However, at very high doses, the acute oxidant effects can be strong enough to induce significant macromolecular damage and/or ROS-dependent apoptosis before the cell can mount a sufficient reductive (antioxidant) response (Zanichelli, Capasso et al. 2012). In fact, such ROS-driven cytotoxic effects are thought to underlie the ability of ITCs to defend cruciferous vegetable plants against insects and microorganisms (Tierens, Thomma et al. 2001).

This phenomenon of a distinctly non-linear relationship between dose and response, in that low-dose exposure exerts beneficial effects, whereas high-dose exposure to the same stimulus exhibits adverse and opposite effects, is known as hormesis. Hormesis is an ancient concept by which various traditional herbal preparations and/or poisons are characterised, in addition to a number of pharmaceutical drugs, micronutrients, macronutrients and phytochemicals. ITCs such as SFN are classic examples of hormetic agents.

Most studies indicate that the doses of ITCs obtainable in humans through the consumption of cruciferous vegetables are far below the cytotoxic threshold, thus firmly within the neutral-to-cytoprotective range. For example, in one study, plasma concentrations of SFN peaked at 0.65 μ M following the consumption of 100g broccoli (150mL broccoli soup) by participants (Janobi, Mithen et al. 2006). In another, the plasma ITC concentration peaked at ~2 μ M following the consumption by participants of processed broccoli sprout preparations, each of which contained 200 μ mol ITCs (predominantly SFN), (Ye, Dinkova-Kostova et al. 2002). Some studies have demonstrated human plasma SFN concentrations to reach 10 μ M following significant broccoli sprout consumption (Hsu, Wong et al. 2011): a concentration which most studies *in vitro* or *in vivo* have indicated to be non-cytotoxic to healthy cells.

The exhibition of Nrf2's reductive (antioxidant) effects in healthy cells is desirable, due to the conferred protection against DNA oxidation-induced mutation. ROS can oxidise guanine bases in DNA to 8-oxoguanine, which may then pair with adenine instead of cytosine, generating a G:C to T:A point mutation. Studies have shown Nrf2-knockout mice to be more susceptible to

the carcinogenic effects of common environmental carcinogens than control mice (Ramos-Gomez, Kwak et al. 2001).

Unfortunately, the cytoprotective effects of Nrf2 do not occur exclusively in non-cancerous cells, but also in cancer cells, and can thereby increase the resistance of the cancer cells to ROS-dependent chemotherapy, and promote their proliferation (Ohta, Iijima et al. 2008). In fact, Nrf2 hyperactivation is observed in certain cancers, and promotes their aggressiveness and chemoresistance by helping to protect the tumour cells against oxidative stress-induced cytotoxicity and/or cytostasis.

The net role that Nrf2 plays in cancer cells is nevertheless highly complex; Nrf2 has been shown to actually exert anti-proliferative effects in certain contexts. For example, the activation of Nrf2 by the miR-200c-5p-mediated repression of Keap1 inhibited the proliferation of human lung cancer cell lines. Nrf2 bound in combination with the polyamine-modulation factor 1 (PMF1) to the polyamine-responsive element (PRE) in the promoter region of the gene coding for the spermidine/spermine N¹-acetyltransferase (SSAT) enzyme, leading to its transcriptional upregulation. The SSAT enzyme breaks down growth-stimulating polyamines, so the Nrf2-mediated upregulation of SSAT inhibited proliferation (Murray-Stewart, Hanigan et al. 2013).

Aside from ITCs, there are various dietary inducers of Nrf2 studied for their chemopreventative potential, such as resveratrol from red grapes, epigallocatechin gallate from green tea, and allicin from garlic. Allicin is demonstrated to be able to inhibit cell proliferation and induce apoptosis at sufficiently high doses. One study surprisingly demonstrated that the anti-proliferative effects of moderate-dose allicin in HCT116 cells were abrogated by the siRNA knockdown of Nrf2 (Bat-Chen, Golan et al. 2010), implying that Nrf2 counterintuitively played a positive role in allicin's anti-proliferative effects. Perhaps this is partially attributable to the reduction of specific proliferation-promoting ROS. Researchers also demonstrated that the livers of mice genetically engineered to have constitutively-active Nrf2 regenerated more slowly following carbon tetrachloride-induced damage, compared to

those of control mice. They also showed that the genes coding for the cyclin-dependent kinase inhibitor p15 and for the pro-apoptotic Bim could both be transcriptionally upregulated by Nrf2, based on chromatin immunoprecipitation sequencing (ChIP-Seq) data (Kohler, Kurinna et al. 2014).

Nrf2 has been shown to induce the Notch-1 protein, which itself has been reported to act in context-dependent oncogenic (Jin, Gong et al. 2013) or tumour suppressive (Sriuranpong, Borges et al. 2001) fashions. Nrf2 also has the potential to induce the expression of interleukin (IL)-17D on the surface of tumour cell lines, thus potentially facilitating natural killer cell-mediated induction of cell death *in vivo* (Saddawi-Konefka, Seelige et al. 2016). Interestingly, hypoxia-inducible factor 1-alpha (HIF-1 α), which is often upregulated in the hypoxic conditions of cancer cells, and is a promoter of angiogenesis, can repress Nrf2 (Loboda, Jozkowicz et al. 2012).

Further complicating the picture of Nrf2's role in cancer is a report suggesting that Nrf2's effects on proliferation depend upon its localisation. Most studies of Nrf2 activity to date have focussed only upon the activity of Nrf2 in the nucleus, since it has often been assumed that cytoplasmically-localised Nrf2 is inactive. However, Lin et al. suggest distinct proliferative and pro-invasive functions for cytoplasmic Nrf2 via its upregulation of the 'proteasome 26S subunit, non-ATPase 4' 26S proteasome subunit, which degrades p53; p53 is a repressor of Crm1 expression, which as mentioned previously exports nuclear Nrf2 to the cytoplasm (Lin, Chang et al. 2016). They also report from clinical tumour studies that the predominance of cytoplasmic rather than nuclear Nrf2 localisation in tumours was associated with worse clinical outcomes (Lin, Chang et al. 2016). One might speculate that the oncogenic effects of inactivating Keap1 mutations observed in certain malignancies are attributable to increased cytoplasmic Nrf2 activity, although contribution from increased nuclear Nrf2 activity would nevertheless be highly probable.

The relationship between redox status and cancer is clearly very complex. SFN has repeatedly been shown to exert tumour suppressive activity towards various cancer cell lines at moderate-to-high doses, by inhibiting

proliferation and promoting cell cycle arrest and apoptosis, which is at least partly attributable to ROS-dependent cytotoxicity. SFN has been shown to induce lipid peroxidation, and thus the accumulation of 4-hydroxynonenal (HNE), which is itself a promoter of cell cycle arrest and apoptosis (Sharma, Sharma et al. 2012).

As previously mentioned, specific ROS function as important cell signalling molecules. For example, intracellular levels of specific ROS such as hydrogen peroxide are increased upon exposure to high glucose conditions, and these ROS promote the secretion of insulin by pancreatic β -cells, in which experimental Nrf2 overexpression can impair the insulin-mediated control of blood glucose levels (Zhang, Pi et al. 2010). Moderate general increases in ROS, i.e. an oxidative shift in redox status, can promote proliferation via the activation of particular extracellular signal-regulated kinase (ERK) and MAPK signalling pathways (Trachootham, Zhou et al. 2006). ROS reportedly interact with the tumour suppressor protein phosphatase and tensin homologue deleted on chromosome 10 (PTEN) in a manner that inhibits its phosphatase activity (Leslie, Bennett et al. 2003), but that perhaps also inhibits its nuclear-to-cytoplasmic export, promoting its phosphatase-independent anti-proliferative effects in the nucleus (Chang, Mulholland et al. 2008).

Excessive oxidative shifts are cytotoxic and promote cell cycle arrest and apoptosis. The therapeutic benefits of many anti-cancer drugs are themselves attributable to ROS- and DNA damage-dependent cytotoxicity towards cancer cells, and their therapeutic window depends largely upon their relative selectivity towards target cells. Experimentally increasing ROS in SiHa (doxorubicin-resistant cervical cancer) cells increased their sensitivity to the anti-cancer drug doxorubicin, whilst experimentally reducing ROS made them more resistant (Filippova, Filippov et al. 2014).

Therefore, on the one hand, Nrf2 expression in cancer cells is often thought to be deleterious because it helps to protect them against oxidative stress-induced death. Nrf2 is overexpressed in several tumours in which it plays an important role in their ability to thrive under stressful extracellular conditions,

and their chemoresistance (Lau, Villeneuve et al. 2008). On the other hand, the ROS-reducing effects of Nrf2 can themselves be anti-proliferative. Particularly in early tumour cells, reducing (antioxidant) effects may tend to inhibit proliferation by reducing the specific ROS-mediated stimulation of ERK and MAPK signalling. Hydrogen peroxide itself is able to activate PKC and thus stimulate anti-apoptotic and proliferative pathways (Prasad and Mishra 2015).

Counterintuitively, the induction of Nrf2 in Calu-6 human lung carcinoma cells increased their sensitivity to the anti-cancer drug BENSpM, perhaps by inducing SSAT and thus increasing the breakdown of endogenous and exogenous proliferative and anti-apoptotic polyamines (Murray-Stewart, Hanigan et al. 2013). Nrf2 has also been reported to inhibit the Smad-dependent migration of certain cell lines (Rachakonda, Sekhar et al. 2010).

1.4 Epigenetics

Epigenetics is the study of how alterations to gene expression levels, as opposed to genetic changes, influence phenotype. The conformation of chromatin (the collective term for the DNA-histone complexes that constitute eukaryotic chromosomes) varies according to the nature and intensity of the DNA-histone interactions, in that tighter interactions tend to repress transcription, whilst looser interactions are more transcription-permitting. Cytosine residues preceding guanine are prone to being methylated by DNA methyltransferases (DNMTs) to 5-methylcytosine, a process generally termed 'DNA methylation'. The nature of DNA-histone interaction is influenced by covalent histone modifications (primarily acetylation and methylation) and DNA methylation. Methylated DNA recruits histone deacetylases (HDACs), and histone methyltransferases (HMTs) via their methyl-CpG-binding domains. HDACs remove acetyl groups from the histones, permitting tighter DNA-histone interactions that inhibit transcription. The methylation of histone H3 at lysine 9 or 27 tends to repress transcription, while that at lysine 4 or 36 actually tends to promote transcription. DNA methylation can also directly inhibit the binding of transcription factors such

as cAMP response element-binding protein, E2F and nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB) (Lao and Grady 2011).



Figure 7: Small regions of the genome that are relatively rich in CpG dinucleotides are called CpG islands. CpG islands have a strong tendency to be found in a methylated state unless located proximal to gene promoters.

CpG dinucleotides are 75% less abundant in the human genome than would be mathematically expected in a random T, G, A, C distribution, which is probably an evolutionary consequence of the liability of 5-methylcytosines to spontaneously deaminate to thymine, which can result in deleterious C to T point mutations.

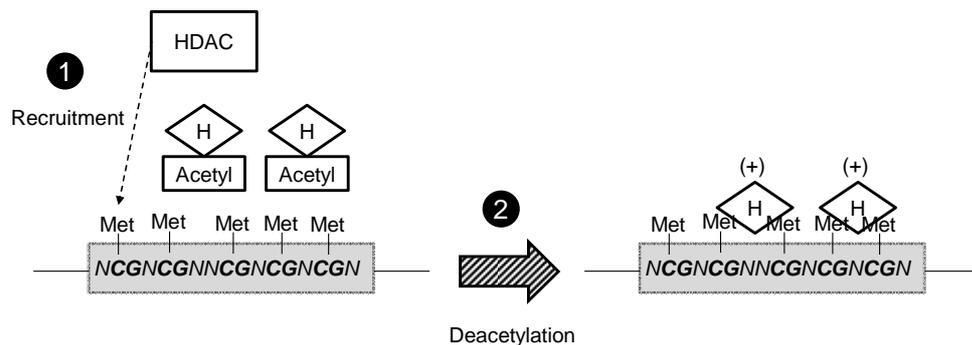


Figure 8: A diagram illustrating the recruitment of HDACs to methylated cytosine residues, and the subsequent repression of gene transcription. (1) Methylated cytosine residues are bound by the methyl-CpG-binding domain of an HDAC, recruiting the HDAC to the site. (2) The HDAC removes acetyl groups, exposing positive charge on the histone tails, which interact with the negatively charged phosphate backbone of DNA. This causes the local chromatin to become more compact, thus sterically inhibiting the binding of transcription factors and machinery.

ITCs are often demonstrated to inhibit the activity of particular HDACs (Nian, Delage et al. 2009), and general HDAC inhibition has itself been demonstrated to be tumour suppressive in a number of contexts via the transcriptional reactivation of TSGs that promote cell cycle arrest and apoptosis. SFN downregulated HDAC3 in the CRC cell line HCT116, in

which basal HDAC3 was demonstrably higher than in several non-cancerous colonic cell lines (Rajendran, Kidane et al. 2013). The activity of HDAC3 can influence genomic stability and cell survival (Rajendran, Kidane et al. 2013).

MicroRNAs (miRNAs) are small, non-coding RNA molecules best characterised as post-transcriptional, translation-inhibiting, regulators of gene expression. One definition of epigenetics that is favoured by some publications includes only heritable, expression-influencing changes to chromatin status, and thus excludes transient modulations of gene expression such as those exerted by miRNAs. However, according to a broader definition of epigenetics that includes all modulation of gene expression in lieu of genetic change, the roles of miRNAs can be included.

1.5 MicroRNAs

As described above, miRNAs are small, non-coding RNA molecules. They are generally 18-25 nucleotides in length and participate in the post-transcriptional control of gene expression via interaction with mRNAs. They originate from various genetic loci such as those in:

- The introns of protein-coding genes.
- The exons of (and adjacent to) protein-coding genes.
- Long exonic clusters that code for multiple miRNAs, also known as miRNA arrays.

1.5.1 Biogenesis

The canonical pathway of miRNA expression in animals is as follows:

1. A genetic locus is transcribed by RNA polymerase II as a 5'-capped and 3'-polyadenylated transcript, typically several hundred nucleotides long, which adopts a secondary structure containing hairpin-loops. This is the primary miRNA transcript (pri-miRNA). A pri-miRNA can contain multiple hairpin-loops motifs that each ultimately give rise to a different miRNA.
2. DiGeorge Syndrome Critical Region 8 (DGCR8) binds to a pri-miRNA hairpin-loop motif and recruits the Drosha enzyme, which excises the

hairpin-loop structure from the pri-miRNA. The excised entity has a 3' 2-nucleotide overhang and is called a precursor miRNA (pre-miRNA).

3. Exportin-5 recognises the 3' 2-nucleotide overhang of a pre-miRNA and transports it from the nucleus into the cytoplasm, where Dicer subsequently recognises the same 3' 2-nucleotide overhang and cleaves open the loop region of the pre-miRNA to generate a linear RNA duplex, with a 3' overhang on both strands.
4. This linear duplex is unwound and one of the strands (typically that with the lower stability at its 5' end) remains bound to Dicer. The miRNA-Dicer complex associates with Argonaute (AGO) proteins, which protect the miRNA against degradation, and enables additional proteins to bind thus leading to the formation of an RNA-induced silencing complex (RISC).
5. The other strand dissociates from Dicer. It was once believed that this 'alternative' strand was typically degraded. However, evidence has since shown that many of these 'alternative' strands actually associate with other AGOs to form other active RISCs.

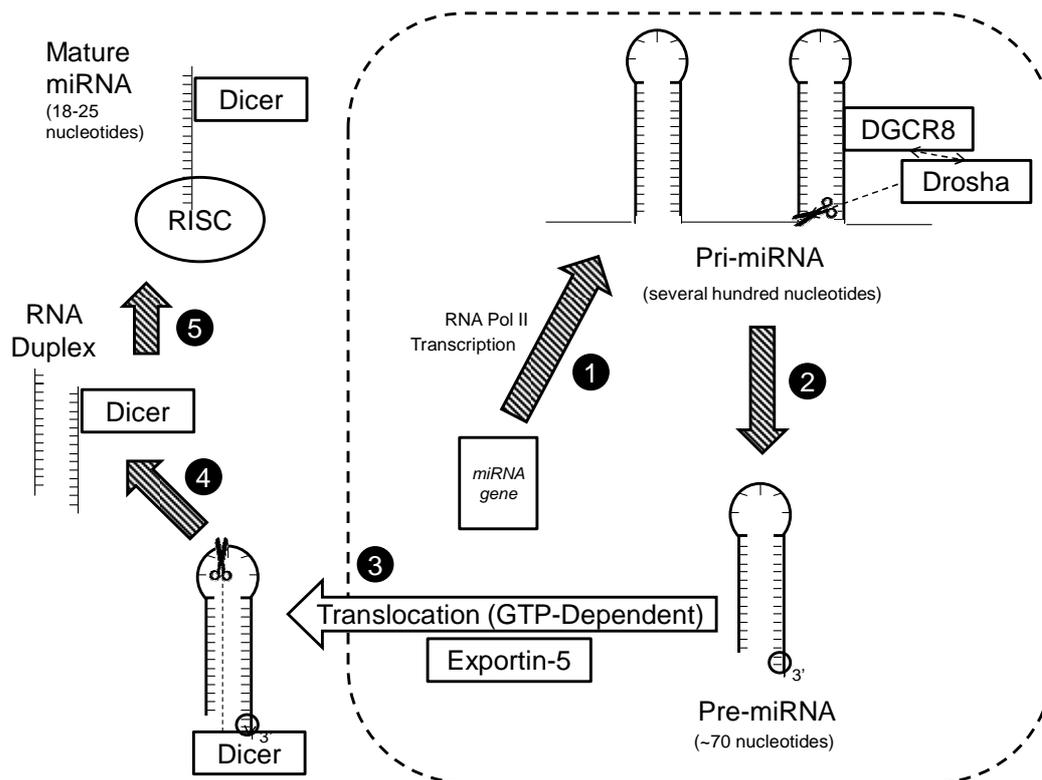


Figure 9: A diagram illustrating the canonical pathway of miRNA expression in animal cells. (1) A genetic locus is transcribed by RNA polymerase II, producing a pri-miRNA, which is several hundred nucleotides long. (2) The pri-miRNA is bound by DGCR8, which recruits Drosha, which cleaves the pri-miRNA into pre-miRNAs that are about 70 nucleotides long and have 2-nucleotide overhangs at their 3' ends. (3) The 3' 2-nucleotide overhangs are recognised by Exportin 5, which uses GTP to transport them from the nucleus to the cytoplasm. (4) Dicer recognises the same 3' overhang and makes a nick in the loop region of the pre-miRNA, generating an imperfectly paired linear RNA duplex, each strand of which bears a 3' 2-nucleotide overhang. (5) The linear RNA duplex is unwound; one strand remains associated with Dicer as the mature miRNA, and becomes part of a RISC upon association with AGOs.

Pri-miRNA transcription and processing are not the only means of pre-miRNA generation. Introns are excised from the early mRNA transcripts of protein-coding genes and released as lariat structures. Some of these lariat structures actually contain miRNA precursor sequences (Westholm and Lai 2011) and can be directly converted to pre-miRNAs by the lariat debranching enzyme (Ldbr), as illustrated in Figure 10.

Pre-miRNAs of a particular secondary structure can be processed independently of Dicer, such that the pre-miRNA itself is directly processed

by AGO2 to form an intermediate that is subsequently uridylated and/or trimmed to generate the mature miRNA (Cifuentes, Xue et al. 2010). Hsa-miR-451a is an example of a miRNA generated in such a fashion (Cifuentes, Xue et al. 2010).

MiRNA biogenesis was classically characterised as a process generating two specific mature miRNAs per pre-miRNA; one from each of the 5' and 3' arms. However, evidence from deep sequencing experiments revealed that multiple isoforms of each mature miRNA are produced, differing slightly according to length and/or sequence, despite originating from the same pre-miRNA. One source of such variation is the fact that Dicer can cleave a pre-miRNA 'imprecisely', thus resulting in 'additional' or 'missing' nucleotides at the 5' or 3' ends. A more upstream source of variation is the fact that Drosha can cleave pri-miRNA transcripts with similar 'imprecision'. Post-transcriptional modifications of pre-miRNAs and/or mature miRNAs by exonucleases or base-modifying enzymes further diversify the pool of isoforms. Variant miRNA isoforms are called isomiRs.

One or two isomiRs of a named miRNA tend to dominate in most contexts. The miRNA sequences listed in miRBase are defined as the canonical forms, and tend to be either the most abundant forms found at the time of accession of the miRNA, or consensus sequences derived from pooled experimental sequencing data. These canonical forms are not necessarily the dominant isoforms in every biological context; there is evidence that the relative abundances of particular isoforms of a given miRNA can vary according to cell and tissue type, and also between the normal and tumour cell counterparts of a tissue (Tan, Chan et al. 2014). Additionally and interestingly, the preferred pre-miRNA cleavage site can be influenced by the association of particular proteins with Dicer. Thus, the abundances of different Dicer-associating proteins can influence the isomiR distribution for a given miRNA (Fukunaga, Han et al. 2012).

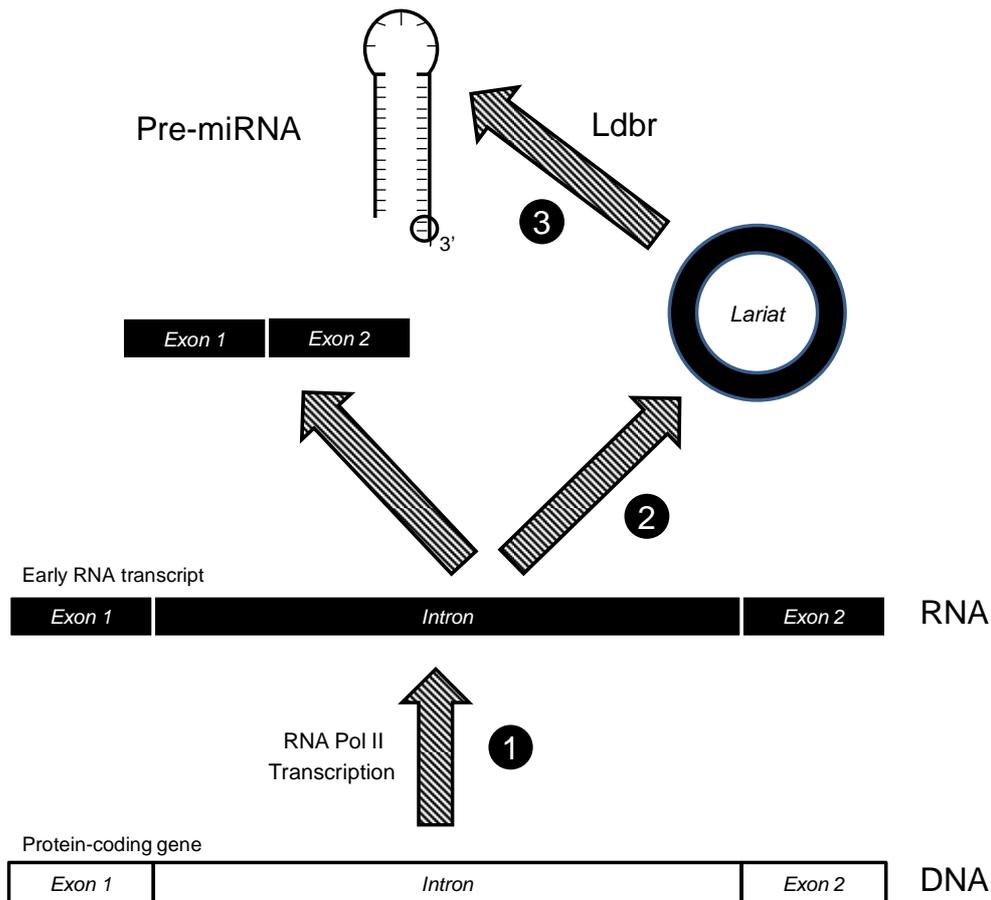


Figure 10: A diagram illustrating the generation of a pre-miRNA from a lariat structure, itself excised from the intronic region of an early mRNA transcript. (1) A genetic sequence is transcribed by RNA polymerase II. (2) The intronic sequence of the RNA transcript is excised in the form of a lariat structure. (3) Ldbr converts the lariat structure into a pre-miRNA.

1.5.2 Canonical Model of Mature MicroRNA Activity

AGO proteins such as AGO1, AGO2, AGO3 and AGO4 have PAZ domains that can interact with the 3' ends of Dicer-bound miRNAs. At the 5' end of a mature miRNA lies a 6-8 nucleotide seed region, which binds with loci in the 3'-untranslated region (3'-UTR) of mRNA transcripts to which the seed region is at least partially complementary in terms of sequence. This interaction tends to facilitate the translational repression and/or degradation of a transcript by the RISC. Perfectly complementary base pairing tends to induce mRNA degradation via the activation of AGO endonuclease activity. Of the four mammalian AGOs: AGO1, AGO2, AGO3 and AGO4, only AGO2 has

thus far been demonstrated to possess such endonuclease activity (Liu, Rivas et al. 2005). However, complementarity is usually partial rather than perfect in animals, and partial complementarity has traditionally been described to repress translation of the mRNA without inducing its degradation (Bartel 2004). However, recent evidence suggests that mRNA degradation occurs even with partial complementarity, more often than previously thought. It is likely that multiple mechanisms contribute to miRNA-mediated repression, the relative significance of each varying according to miRNA, target, cell type and environment (Hammell 2008).

P-bodies and stress granules (SGs) are cytoplasmic loci involved in mRNA degradation and translational repression respectively, and the formation of both loci can be induced by miRNAs. MiRNAs have been shown to directly induce the formation of P-bodies in mammalian cells, and the knockdown of Drosha/DGCR8 (Schmittgen 2008) or direct inhibition of RISC formation has been shown to inhibit P-body formation (Zhao and Liu 2009). Reciprocally, experimentally blocking the formation of P-bodies has been shown to impair miRNA-mediated translational repression (Zhao and Liu 2009). RISCs have been demonstrated to associate with SGs, suggesting that both SGs and P-bodies may play important roles in miRNA-mediated repression (Zhao and Liu 2009). AGOs can directly interact with SGs in an miRNA-dependent manner (Leung, Calabrese et al. 2006). The hydroxylation and/or phosphorylation of AGO2 at particular sites can facilitate its association with P-bodies (Bonfrate, Altomare et al. 2013).

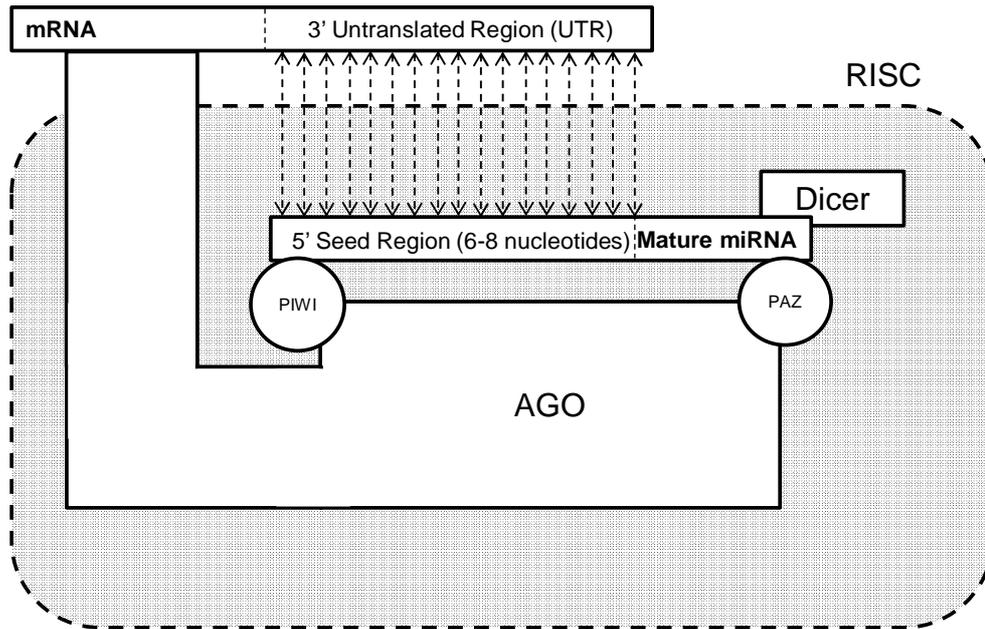


Figure 11: A diagram illustrating the interaction between a mature miRNA, an AGO protein, and a target mRNA. The AGO interacts with the miRNA and at the 5' and 3' ends, via its PIWI and PAZ domains respectively. A 5' seed region of the miRNA base pairs with part of the 3' untranslated region of a target mRNA, to which it is partially or fully complementary. Subsequently, the AGO can interact with the target mRNA to inhibit its translation.

Any given miRNA has the potential to target a multitude of mRNAs, which are each themselves susceptible to regulation by various miRNAs. MiRNAs are known to regulate the expression of at least 30% of protein-coding genes in humans (Lewis, Burge et al. 2005), thereby playing crucial roles in development and health. Amongst such miRNA-regulated genes are many oncogenes and TSGs i.e. those that regulate apoptosis, cell cycle progression, DNA repair, senescence, inflammation and redox status (Calin and Croce 2006).

1.5.3 IsomiRs

As mentioned above, the cleavage of pre-miRNAs by Dicer does not solely generate a specific pair of 5' and 3' products, but rather a distribution of isoforms called isomiRs, that vary from the canonical forms by having 'extra' or 'missing' nucleotides at the 5' or 3' ends, or in rare cases, base pair

'substitutions'. In some cases, such variation might actually alter the seed regions and thus change mRNA-targeting activities. As previously mentioned, any given mature miRNA has the potential to target multiple mRNA transcripts. It is interesting to consider the degree of target overlap between particular isomiRs and the implications of such. Particular mRNAs that are targeted by multiple isomiRs of a given miRNA may be more strongly regulated than those mRNAs that are targeted by the canonical isoform alone. It has been hypothesised that this phenomenon may have evolved due to the intensification of the desirable regulation of specific mRNA targets, whilst diffusing effects on undesirable targets (Cloonan, Wani et al. 2011). In this hypothesis, where simply doubling the level of a specific canonical form would equally intensify the regulation of both its desirable and undesirable targets, alternatively co-expressing a different isoform at an equivalent level - an isoform that desirably targets a specific mRNA like the canonical form, but that unlike the canonical form does not deleteriously target one or several other mRNAs – may prove to be advantageous. This hypothesis is illustrated in Figure 12.

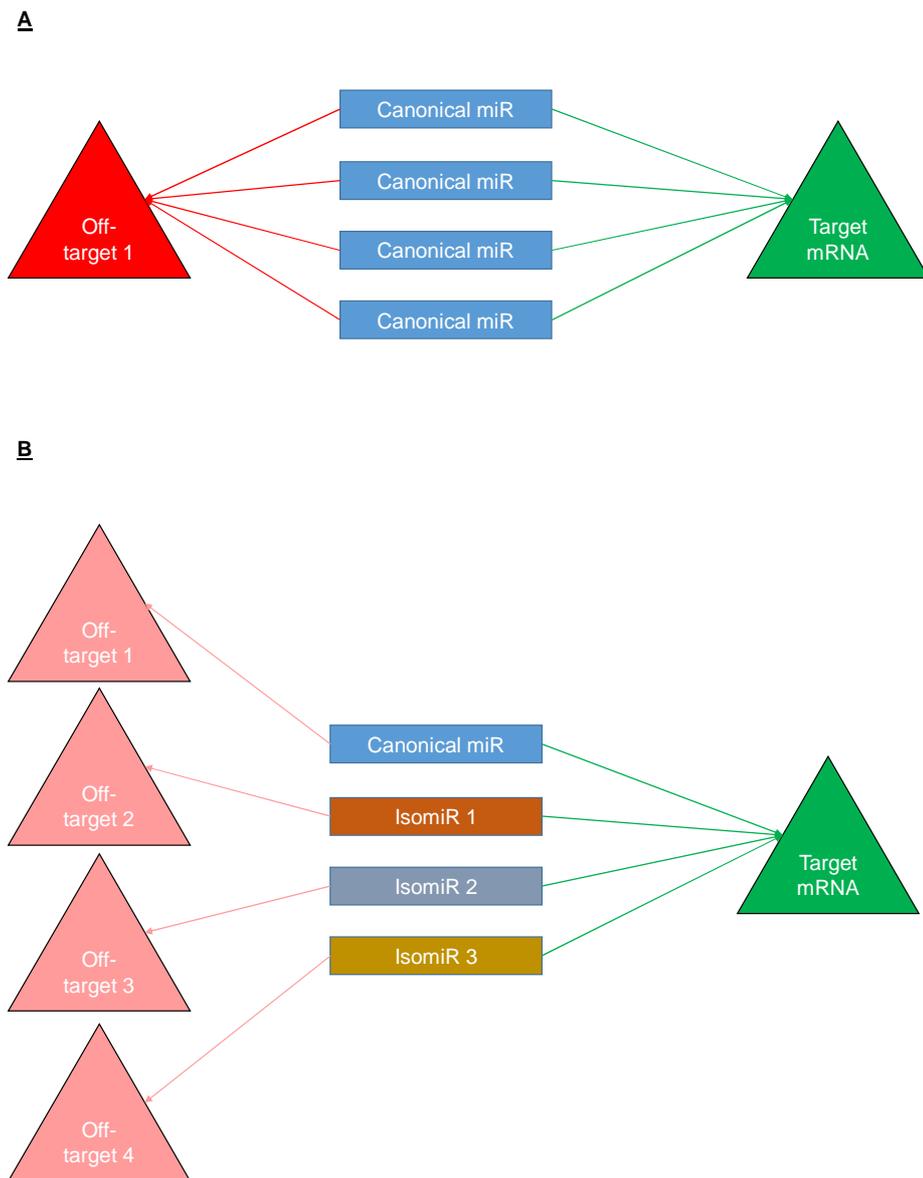


Figure 12: An illustration of how the generation of isomiRs may help to intensify the desirable regulation of a specific target, without the undesirable intensification of regulation of another target. (A) The amplification of the canonical miRNA increases the regulation of the desired mRNA target, but also undesirably intensifies the regulation of an 'off-target' mRNA. (B) The generated isomiRs all contribute to the desirable regulation of the target mRNA, but do not necessarily share the same 'off-target' mRNAs, thus potentially undesirable effects on 'off-target' mRNAs are diluted.

The extent of target overlap between isoforms is dependent upon the site at which they vary; 5' variants naturally tend to be less conserved in their

targeting since it is at the 5' end that seed regions are located (Tan, Chan et al. 2014).

1.5.4 Alternative Effects and Mechanisms of Activity

MiRNAs are mostly characterised as translational repressors. However, miRNAs can actually upregulate the translation of their targets in certain contexts. For example:

- MiR-369-3p represses the translation of its target - tumour necrosis factor- α (TNF- α) - under basal conditions, but actually upregulates its translation under serum-starved conditions (Vasudevan, Tong et al. 2007). The promotion of TNF- α translation is facilitated by the association of miR-369-3p specifically with AGO2, which itself associates with fragile X mental retardation-related protein 1 (FXR1), a protein essential for the association of miR-369-3p with AGO2 (Vasudevan, Tong et al. 2007). Hypothetically, conditions of serum starvation could alter the solubility and subcellular localisation of AGO2-FXR1 complexes in a manner facilitating their association with the miRNA (Vasudevan, Tong et al. 2007). AU-rich elements present at the 3' end of a transcript tend to increase the transcript's susceptibility to degradation (Chen and Shyu 1995). The TNF- α transcript has such an AU-rich element at its 3' end. However, this AU-rich element actually facilitates the miRNA-mediated upregulation of TNF- α expression (Vasudevan, Tong et al. 2007).
- MiR-10a-5p upregulated the translation of ribosomal proteins in E14 ES mouse embryonic stem cells under amino acid starvation, upon binding to 5'-UTRs of ribosomal protein mRNAs at loci immediately downstream of their 5'-terminal oligopyrimidine (5'-TOP) motifs (Ørom, Nielsen et al. 2008).

The potential for miRNAs to differentially affect their targets according to nutrient status could have implications for their roles in carcinogenesis at different stages, and their interactions with cancer therapies. It has been postulated that miRNAs generally repress the translation of their targets in

dividing cells, whilst instead promoting translation in quiescent cells (Zhao and Liu 2009).

Certain miRNAs can function non-canonically as decoys for RNA-binding proteins such as the hnRNP E2 protein: an RNA-binding protein that represses the transcription of CCAAT/enhancer-binding protein alpha (CEBP α), which is a myeloid differentiation factor that acts as a tumour suppressor in chronic myeloid leukaemia (CML). By binding to hnRNP E2, miR-328-3p inhibits the hnRNP-mediated repression of the CEBP α gene (Wilczynska and Bushell 2015), thus bolstering CEBP α -mediated tumour suppression. MiR-328-3p has been shown to be underexpressed in CML patients (Wilczynska and Bushell 2015).

1.6 Links between ITCs, MicroRNAs, Biochemical Pathways and Cancer

A Google Scholar search on 17/02/17 for 'intitle:microrna|mirna|micrornas|mirnas|"mir" intitle:colorectal|colon|rectal|bowel intitle:cancer|tumour|tumor|cancers|tumours|tumors|carcinogenesis|tumorigenesis' returned reports of 144 miRNAs as functionally involved in CRC pathogenesis. 85 of these were reported as tumour-suppressive, 45 as oncogenic, and 14 as ambiguous.

1.6.1 Reported Modulation of MicroRNAs by Isothiocyanates

- Benzyl isothiocyanate (BITC) treatment was shown to downregulate the oncogenic miR-375 and to upregulate the tumour suppressor miR-221-3p in pancreatic cancer (Basu, Alder et al. 2011), thus opposing the changes in expression observed in cancerous vs. non-cancerous cells.
- Phenethyl isothiocyanate (PEITC) was demonstrated to inhibit environmental cigarette smoke-induced oncogenic miRNA modulation in a neonatal mouse model (Izzotti, Larghero et al. 2010).
- ITCs were demonstrated to upregulate the tumour suppressing miR-23b-3p and miR-27b-3p, and to downregulate the oncogenic miR-155-

5p, in the NCM460 and NCM356 CRC cell lines (Slaby, Sachlova et al. 2013).

- Slaby et al. screened 754 human miRNAs for 48 h SFN-induced modulation by qPCR arrays in two non-cancerous colorectal cell lines – NCM460 and NCM356 - using TaqMan Low Density Arrays. The data suggested that a number of miRNAs may be modulated by SFN, as summarised in Table 1.

Table 1: A summary of miRNAs reported to be altered by 48 h SFN treatment in NCM356 or NCM460 cells according to microRNA low-density array RT-qPCR experiments (Slaby, Sachlova et al. 2013), along with potential functions reported in the literature.

Regulation by SFN	MicroRNA: hsa-miR-	Reported Functions
Up	9-5p	Was shown to promote motility, invasiveness and β -catenin signalling by targeting E-cadherin in HMLE and SUM149 mammary epithelial cells, (Ma, Young et al. 2010), and is itself activated by MYC. However, its expression was positively associated with recurrence-free survival time in clear cell renal carcinoma patients, while its hypermethylation was negatively associated with such (Hildebrandt, Gu et al. 2010).
	9-3p	Found to inhibit the growth of MDA-MB-231 breast cancer cells by targeting β_1 integrin (Zawistowski, Nakamura et al. 2013).
	23b-3p	Demonstrated to inhibit human hepatocellular carcinoma cell migration and proliferation by targeting and inhibiting urokinase-type plasminogen activator and c-met (Salvi, Sabelli et al. 2009).
	27b-5p	Predicted to target prospero homeobox protein 1 (DIANA-Lab), which is a promoter of colorectal adenocarcinoma progression (Petrova, Nykänen et al. 2008).
	27b-3p	Shown to repress the translation of matrix metalloproteinase 13 in human osteoarthritis chondrocytes (Akhtar, Rasheed et al. 2010).

(Table 1 Continued)

Regulation by SFN	MicroRNA: hsa-miR-	Reported Functions
	30a-3p	Unclear, although 30a-5p was shown to inhibit the growth and promote the cell cycle arrest of HCT116 CRC cells, by targeting denticleless protein homolog in HCT116 CRC cells (Baraniskin, Birkenkamp-Demtroder et al. 2012).
	135b-3p	Predicted to target several zinc finger proteins (DIANA-Lab); some of which the potential to promote cisplatin resistance has been demonstrated in several cell lines (Duan, Choy et al. 2009).
	145-5p	Found to inhibit the expression of the oncoprotein c-Myc in breast cancer and CRC cells (Sachdeva, Zhu et al. 2009), and to inhibit the growth of bladder cancer cells (Chiyomaru, Enokida et al. 2010).
	146a-5p	Shown to inhibit the proliferation of hormone-refractory prostate carcinomas, in which it tends to be underexpressed, and to target ROCK1, which is putatively involved in transformation (Lin, Chiang et al. 2008). Also demonstrated to reduce the invasiveness of pancreatic cancer cells (Li, VandenBoom et al. 2010).
	342-3p	Found to sensitise breast cancer cell lines to tamoxifen-induced apoptosis, and has a tendency to be underexpressed in tamoxifen-resistant breast cancer cells (Cittelly, Das et al. 2010).
	372(-?p)	'MiR-372' (not stated whether -5p or -3p) is reportedly oncogenic in the human gastric adenocarcinoma cell line, AGS (Cho, Shin et al. 2009).
	486-5p	Demonstrated to induce senescence and inhibit proliferation of human adipose tissue-derived mesenchymal stem cells (Kim, Hwang et al. 2012).
	505-3p	Found to induce apoptosis in NCF7-ADR human breast cancer cells (Yamamoto, Yoshioka et al. 2011).

(Table 1 Continued)

Regulation by SFN	MicroRNA: hsa-miR-	Reported Functions
	629-5p	Shown to be able to suppress lung and breast carcinogenesis by targeting Nibrin (Li and Lu 2013).
	758-3p	Predicted to target <i>BACH2</i> (DIANA-Lab) – a promoter of oxidative stress-induced apoptosis (Spira, Beane et al. 2007).
Down	106a-3p	Found to stimulate the proliferation of HCT116 CRC cells (Wu, Tsai et al. 2011), and predicted to target G protein-coupled receptor kinase 5 (DIANA-Lab).
	155-5p	Was able to promote the growth of several human breast cancer cell lines by targeting <i>Suppressor of Cytokine Signalling 1</i> , and to promote tumour progression <i>in vivo</i> in mouse models, partly by enhancing janus kinase (JAK)-signal transducer and activator of transcription 3 (STAT) signalling (Jiang, Zhang et al. 2010).
	633-3p	May target TGF- β 1, which has dual roles in cancer.

1.6.1.1 Examples of MicroRNA Modulation by ITC-Influenced Factors

ITCs have the potential to modulate miRNA expression by various means, for example:

- By inhibiting the activity of DNMTs and HDACs (Myzak, Hardin et al. 2006), thus transcriptionally upregulating genetic loci that give rise to miRNAs.
 - HDAC3 repressed the expression of the pro-apoptotic miR-15a-5p/16-1-5p cluster in mantle cell lymphoma, members of which target the anti-apoptotic protein Bcl-2 (Zhang, Chen et al. 2012). SFN and other ITCs (15 μ M) were shown to inhibit HDAC3 activity in HCT116 cells, to promote DNA damage and autophagy, and to inhibit proliferation (Rajendran, Kidane et al. 2013). By contrast, similar treatment of the non-cancerous

colonic cell line CCD-841 did not produce similar effects (Rajendran, Kidane et al. 2013). Perhaps the steady-state redox status of HCT116 cells is more oxidative than that of CCD-841 cells, making them more sensitive to the acute pro-oxidant effects of ITCs.

- By modulating redox status, to which certain microRNAs are sensitive (Xu, Zhang et al. 2012).
 - Oxidative stress has been shown to modulate miRNA expression in both mouse hippocampal neuronal cells (Xu, Zhang et al. 2012), and normal human fibroblasts (Simone, Soule et al. 2009).
 - Oxidative stress has been demonstrated to inhibit Dicer activity in JAR trophoblast cells, thus globally downregulating the generation of mature miRNAs (Wiesen and Tomasi 2009).
 - DGCR8, which is involved in the cleavage of pri-miRNAs to pre-miRNAs, was shown to be inhibited by oxidative stress (Cheng, Ku et al. 2013). The HO-1 enzyme catabolises ferric haem - a promoter of DGCR8 activity - so the oxidative stress-induced repression of DGCR8 activity might be due to Nrf2-HO-1-mediated ferric haem catabolism.
 - Oxidative stress was demonstrated to activate Drosha by the induction of the redox-sensitive glycogen synthase kinase 3 β (GSK3 β), which phosphorylates Drosha at Ser300 and Ser302; such modification promotes the translocation of Drosha to the nucleus where it is able to cleave pri-miRNAs to form pre-miRNAs (Cheng, Ku et al. 2013).
 - Endoplasmic reticulum (ER) stress is a typical consequence of oxidative stress. ER stress induces an ER transmembrane kinase—endoribonuclease (RNase) called inositol-requiring enzyme 1 α (IRE1 α), which has been shown to be able to degrade the pre-miRNAs giving rise to miR-17-5p, miR-34a-5p, miR-96-5p and miR-125b-5p.

- Oxidative stress was able to induce the ADP-ribosylation of AGO2 in liver cancer HeLa cells, which inhibited AGO2's ability to repress mRNA translation (Leung, Vyas et al. 2011).
- By inhibiting inflammatory signalling.
 - A pro-inflammatory medium (lipopolysaccharide (LPS)-stimulated macrophage-condition medium containing the cytokines TNF, IL-6, IL-8 and IL-1 β) was shown to upregulate miR-155-5p in MDA-MB-231, BC-453 and T47D breast cancer cells. MiR-155-5p can target anti-mutagenic genes and promote proliferation (Tili, Michaille et al. 2011).
 - Treatment of HCT15 and HCT116 CRC cells with a similar inflammatory medium increased the expression of miR-146a-5p (Tili, Michaille et al. 2011).
 - The transcription factor NF- κ B mediates inflammatory signalling, and was shown to upregulate miR-34a-5p in EC109 oesophageal squamous cancer cells (Li, Wang et al. 2012).
- By influencing the expression of proteins involved in miRNA biogenesis, such as RNA polymerase II, Dicer, Drosha, DGCR8, Exportin 5, Ldb1 and AGO.
- By inducing Nrf2.
 - Some genetic loci that give rise to miRNAs actually have an ARE in their 5' flanking regions, through which they may be liable to Nrf2-mediated transcriptional regulation.
 - ChIP-Seq identified many potential Nrf2-binding sites in SFN-treated lymphoid cells that were proximal to miR-1, miR-29b-3p, miR-32-5p, miR-181c-5p, miR-193b-3p, miR-200c-3p, miR-206, miR-365a-3p, miR-550-5p, miR-592, miR-617, and miR-1207-5p (Shah, Rushworth et al. 2013). MiR-29b-3p expression was shown to be downregulated as a consequence of Nrf2 binding (Shah, Rushworth et al. 2013).
 - Nrf2 was shown to induce the transcription of the pre-miRNAs hsa-mir-125b-1 (giving rise to miR-125b-5p and miR-125b-3p) and mir-29b-1 (giving rise to miR-29b-5p and miR-29b-3p) in

acute myeloid leukaemia cells (Kurinna and Werner 2015).

Interestingly, miR-125b-5p has been shown to translationally repress the tumour suppressing p53 in human CRC cells (Nishida, Yokobori et al. 2011).

- Nrf2 was demonstrated to induce the expression of HDAC4 in A549 human lung carcinoma cells, consequently leading to the transcriptional repression of the genetic loci giving rise to miR-1-3p and miR-206, thus an oncogenic effect (Singh, Happel et al. 2013).
- Less directly, the modulation of redox status and inflammatory signalling that results from Nrf2 activation can influence the expression of various miRNAs

Further complicating the picture is the fact that HDACs/DNMTs, inflammation, Nrf2 and redox status are each able to interact with one another, as described in Table 2.

Table 2: A table illustrating some of the interactions between HDACs/DNMTs, inflammation, Nrf2 and redox status.

	HDACs/DNMTs	Inf.	Nrf2	Redox Status
HDACs/ DNMTs		The HDAC inhibitor ITF2357 was shown to inhibit inflammatory cytokine expression in LPS-stimulated peripheral blood mononuclear cells (Leoni, Fossati et al. 2005).	The forced downregulation of HDAC2 was shown to inhibit the stabilisation and activity of Nrf2 in BEAS-2B bronchial epithelial cells (Mercado, Thimmulappa et al. 2011).	The ROS-generating <i>DUOX</i> NADPH oxidases have a demonstrable tendency to be hyper-methylated in several lung cancer cell lines (Luxen, Belinsky et al. 2008).

(Table 2 continued)

	HDACs/DNMTs	Inf.	Nrf2	Redox Status
Inf.	The processes of inflammation tend to lower extracellular pH (Bäckdahl, Bushell et al. 2009). Low pH can inhibit the acetylation of histones H3 and H4 (McBrian, Behbahan et al. 2013).			Inflammatory stimuli were shown to induce the superoxide anion-generator, NADPH-oxidase, in phagocytes. (Locatelli, Canaud et al. 2003).
Nrf2	Nrf2 deletion was demonstrated to reduce HDAC2 activity in the mouse lung (Adenuga, Caito et al. 2010). Perhaps this was due to increased ROS, which themselves have been shown to inhibit HDAC2 in BEAS-2B cells (Ito, Hanazawa et al. 2004).	The overexpression of Nrf2 was shown to downregulate several inflammatory proteins in human aortic endothelial cells (Chen, Dodd et al. 2006).		As is well-established, Nrf2 activates reducing (antioxidant) enzymes, such as TrxR-1, GST- α , HO-1, NQO1 and UGT, upon binding to the <i>ARE</i> in their gene promoter regions.
Redox Status	ROS were demonstrated to inhibit HDACs and activate histone acetyltransferases (Rahman, Marwick et al. 2004). Hydrogen peroxide was shown to inhibit HDAC2 in BEAS-2B cells (Ito, Hanazawa et al.	Oxidative stress was shown to promote I κ B degradation, thus derepressing the pro-inflammatory transcription factor NF- κ B (Poynter and Daynes 1998).	As is well-established, ROS induce Nrf2 by various means such as by modifying certain cysteine residues of Keap1 in a manner that weakens Keap1-Nrf2 interactions, thus enabling	

(Table 2 continued)

	HDACs/DNMTs	Inf.	Nrf2	Redox Status
	2004). Increased ROS inevitably stimulates GSH synthesis, a process that inhibits the regeneration of methionine and folate - donors of methyl groups that are required for DNMT-induced methylation (Hitchler and Domann 2007).		Nrf2 to translocate to the nucleus.	

1.6.2 Complexities in the Regulation of MicroRNA Abundance and Activity

The known mechanisms by which miRNAs and miRNA-mediated effects are able to be modulated extend beyond those described above, and will most likely continue to increase with time.

1.6.2.1 Interfering RNAs

Competing endogenous RNAs (ceRNAs) are transcripts that are able to inhibit the mRNA-targeting activity of miRNAs by acting as miRNA binding decoys (Wilczynska and Bushell 2015). An example of a ceRNA is the high-mobility group AT-hook 2 (HMGA2) mRNA, which can act oncogenically in lung cancer independently of its translational protein product; it serves as a decoy for miRNAs of the let-7 family, thereby inhibiting their binding to - and transcriptional repression of - mRNAs for oncogenes such as TGF β receptor III (Wilczynska and Bushell 2015).

There exist pseudogene homologues of protein-coding oncogenes, whose transcripts similarly act as decoys for miRNAs. One such example is phosphatase and tensin homolog (mutated in multiple advanced cancers 1) (PTENP1), whose mRNA transcript has been demonstrated to inhibit the binding of miRNAs to - and thus translational repression of - the protein-coding PTEN transcript (Wilczynska and Bushell 2015).

Several long non-coding RNAs (lncRNAs) also serve as miRNA decoys, such as the *Highly Up-regulated in Liver Cancer* lncRNA which serves as such for miR-372-3p in liver cancer cells (Wilczynska and Bushell 2015). Circular RNAs (circRNAs) are a distinctive species formed by the covalent linkage of the 5' and 3' ends of exonic transcripts by the spliceosome, whose circularity enables them to resist linearity-dependent RNA degradation. A circRNA called 'circular RNA sponge for miR-7' was demonstrated to act as a functionally significant decoy for miR-7-5p in brain tissues (Wilczynska and Bushell 2015).

1.6.2.2 Interfering Proteins

The processing and stability of pri-miRNAs and pre-miRNAs are influenced by interactions with various biomolecules. For example, the interaction of Receptor-regulated mothers against decapentaplegic (R-Smad) proteins with pri-miR-21 has been demonstrated to enhance the Drosha-mediated processing of the latter (producing pre-mir-21, which gives rise to miR-21-5p and miR-21-3p) (Hata and Lieberman 2015). Interestingly, p53 was shown to act as a co-factor of the Drosha microprocessor complex to enhance pri-miRNA processing, and thereby upregulate the pro-apoptotic miR-203a-3p (Hata and Lieberman 2015). The KH-type splicing protein (KSRP) was shown to promote the Drosha-dependent processing of pri-mir-21, pri-mir-125, and the let-7 pri-miRNAs (Hata and Lieberman 2015). The activity of KSRP was demonstrated to be enhanced by its own phosphorylation, which is induced in response to DNA damage and by the phosphoinositide 3-kinase-RAC-alpha serine/threonine-protein kinase (PI3K-AKT) pathway (Hata and Lieberman 2015). KSRP was also shown to promote the exportin 5-mediated nuclear-to-cytoplasmic transport of pre-miRNAs, and their subsequent processing by Dicer. The p72 RNA helicase, whose own activity

can be inhibited by nuclear Yes-associated protein (YAP) via sequestration, was shown to facilitate the Drosha-dependent processing of some pri-miRNAs (Hata and Lieberman 2015). Also, pleckstrin homology domain-containing family A member 7 (PLEKHA7) was demonstrated to promote the processing of pri-mir-30b and pre-mir-30b in the apical regions of cell junctions, by interacting with E-cadherin and recruiting Drosha and DGCR8 (Kourtidis, Ngok et al. 2015). In addition, the endonuclease NCP-induced protein 1 was shown to reduce the stability of pre-mir-135b and pre-mir-146a upon binding to their terminal loops (Hata and Lieberman 2015).

At the mature miRNA level, several miRNAs including miR-221-3p, were shown to be selectively degraded by the interferon (IFN)-inducible exonuclease polyribonucleotide nucleotidyltransferase 1 (PNPT1) in melanoma cells (Hata and Lieberman 2015).

1.6.2.3 Post-transcriptional MicroRNA Modifications

Pri-miRNAs, pre-miRNAs and mature miRNAs are all potentially susceptible to activity-influencing post-transcriptional modifications. Such modifications of mature miRNAs or their precursors can affect their stability, for example:

- The 3'-uridylation of let-7 family pre-miRNAs was shown to promote their degradation by DIS3 mitotic control homolog (*S. cerevisiae*)-like 2 (Hata and Lieberman 2015).
- In human breast cancer cells, the demethylation of the 5' monophosphate of pre-mir-145 by the RNA methyltransferase 'BCDIN3 domain containing RNA methyltransferase' inhibited its Dicer-dependent processing (Hata and Lieberman 2015).
- The 3'-adenylation of miR-122-5p by the RNA nucleotidyl transferase 'PAP-associated domain containing 4' was shown to increase its stability, while the 3'-uridylation of miR-26a-5p by an RNA nucleotidyl transferase called terminal uridylyltransferase 4 repressed its mRNA transcript-binding activity (Wyman, Knouf et al. 2011).

Some post-transcriptional modifications of miRNAs generate variants that do not map in sequence with the genetic loci from which they originate, such as non-templated additions (NTAs). It has been reported that only a minority of

human miRNAs are susceptible to 3' NTAs, but the susceptibility of a given miRNA may be independent of cell and tissue type and/or disease status (Wyman, Knouf et al. 2011). The most frequently observed forms of 3' NTA are mono-adenylation and mono-uridylation, but the relative dominance of each form may vary between species (Wyman, Knouf et al. 2011).

Adenosine deaminases (ADARs) are enzymes that tend to exist in either monomeric or dimeric forms. In the latter, they modify adenosine residues to inosine in RNAs such as pri- and mature miRNAs, which can have complex effects on miRNA regulation. For example, the editing of pri-mir-142 by ADAR1 and ADAR2 was shown to inhibit its Drosha-dependent processing and to promote its degradation. Additionally, ADAR1 appeared to generally inhibit Drosha-dependent pri-miRNA processing by competitively inhibiting DGCR8 recruitment (Hata and Lieberman 2015). ADAR-mediated editing of miRNAs has the potential to change their mRNA-targeting properties, especially if occurring in seed regions (Nielsen, Goodall et al. 2012). Interestingly, monomeric ADARs were found to have a tendency to promote both the Dicer-dependent processing of pre-miRNAs and the loading of mature miRNAs into RISCs.

As mentioned previously, Drosha and Dicer do not always cleave precisely at specific loci of pri-miRNAs and pre-miRNAs, so isomiRs that vary by length but that still map to their genetic loci of origin are generated; this is called templated variation (Nielsen, Goodall et al. 2012). Templated variation can also arise from the action of exonucleases on miRNAs (Nielsen, Goodall et al. 2012).

1.6.2.4 General Regulation

- Posttranslational modifications of the RNA processing enzymes Drosha, DGCR8 and Dicer, have the potential to influence their stability, localisation and activity (Hata and Lieberman 2015).
- The sequence of a given mature miRNA influences its innate susceptibility to degradation under certain conditions.

- The miR-15/16 family of miRNAs have been demonstrated to undergo rapid degradation as cells re-enter the cell cycle from G0 arrest (Wilczynska and Bushell 2015).

1.6.3 Interactions between MicroRNAs and Nrf2

MicroRNAs can both promote and inhibit Nrf2 activity. For example, the miR-200-3p/141-3p family were shown to promote Nrf2 activity by repressing Keap1 in hepatic stellate cells. Interestingly, SFN (5 μ M) was shown to upregulate miR-200c-3p in T24 cells (Shan, Zhang et al. 2013), thus one could speculate that miR-200c-3p induction contributes to the SFN-mediated induction of Nrf2. MiR-153-3p was shown to repress Nrf2 and inhibit GSH synthesis in human neuroblastoma cells, while miR-34a-5p and miR-93-5p were shown to repress Nrf2 in the rat liver (Kurinna and Werner 2015). MiR-28-5p was also shown to repress Nrf2 in breast cancer cells, and miR-507, miR-634, miR-450a-5p and miR-129-5p to do so in oesophageal squamous cell carcinoma cells, in which the downregulation of these miRNAs was observed to contribute to oncogenesis (Kurinna and Werner 2015).

The zinc-finger enhancer binding (ZEB) transcription factors can promote epithelial-mesenchymal transition (EMT): an important step of oncogenic transformation. The miR-200 family of miRNAs have been demonstrated to repress the ZEB genes (Burk, Schubert et al. 2008), and their experimental overexpression to do so in MDCK cells (Gregory, Bert et al. 2008). ZEB-1 was shown to downregulate E-cadherin in MCF10A mammary epithelial cells (Chua, Bhat-Nakshatri et al. 2006); E-cadherin plays a role in tumour suppression, by maintaining physical cell-cell contacts thus repressing cell migration, and also via the induction of Twist-activated tumour-suppressive signalling (Onder, Gupta et al. 2008).

1.6.4 MicroRNAs in Colorectal Cancer

As previously described, many miRNAs have potential roles in CRC pathogenesis.

1.6.4.1 MicroRNA Interactions with Vogelstein Model

Pathogenesis

Several miRNAs are able to interact with elements of the canonical Vogelstein CRC pathogenesis model (previously summarised in Figure 2):

1. The loss of *APC* leads to the activation of β -catenin, which promotes stem cell renewal and proliferation, facilitating the development of an ACF. Two miRNAs often overexpressed in CRC - miR-135a-5p and miR-135b-5p - can repress *APC*, and thus potentially promote this step of carcinogenesis (Slaby, Svoboda et al. 2009), whilst miR-17-5p can promote β -catenin activity (Amirkhah, Schmitz et al. 2015). Meanwhile, miR-320a can inhibit tumour growth by repressing β -catenin (Amirkhah, Schmitz et al. 2015).
2. Increased *KRAS* activity typically arises from hyperactivating mutations and increases proliferation, facilitating the development of the ACF into an early adenoma. The let-7 family of miRNAs, along with miR-18a-5p, miR-143-3p (Slaby, Svoboda et al. 2009) and miR-18a-3p (Amirkhah, Schmitz et al. 2015) have all been shown to repress *KRAS*.
3. The losses of the pro-apoptotic gene *DCC* and the TGF β -driven TSGs *SMAD2* and *SMAD4* facilitate the progression of the adenoma to its late stage. The miR-130a/301a/454-3p family and miR-224-5p have been shown to repress *SMAD4* itself, and miR-106a-5p and miR-21-5p to repress the TGF β receptor II (70/80kDa) (Amirkhah, Schmitz et al. 2015).
4. The mutational inactivation of the *TP53* gene is canonically considered to be the final critical step that facilitates the progression of a late adenoma to a metastatic cancer. Some of p53's tumour-suppressive effects are mediated by the induction of the miR-34a-c family. MiR-34a-5p represses the HDAC sirtuin 1 (*SIRT1*) and is induced in response to DNA damage in a p53-dependent manner (Amirkhah, Schmitz et al. 2015). The miR-34a-c family of miRNAs is frequently deleted or hypermethylated in CRC (Amirkhah, Schmitz et al. 2015). MiR-34a-5p-based therapy is currently undergoing phase I

clinical trials for the treatment of liver cancer (Amirkhah, Schmitz et al. 2015).

1.6.4.2 Other Potential Interactions between MicroRNAs and Colorectal Cancer

There are various other examples of potential miRNA-CRC interaction. For example:

- MiR-145-5p can repress insulin-like growth factor-1 receptor (IGF-1R) and insulin receptor substrate (IRS-1), both of which promote proliferation and inhibit apoptosis in response to growth factors such as insulin-like growth factor 1 (IGF-1) (Slaby, Svoboda et al. 2009).
- MiR-126-3p can inhibit the oncogenic PI3K signalling pathway by repressing one of its promoters - p85 β (Slaby, Svoboda et al. 2009).
- MiR-21-5p can conversely promote PI3K signalling by repressing the PI3K signalling-inhibitor PTEN (Slaby, Svoboda et al. 2009).
- Interestingly, the oncogene *c-Myc* can induce miRNAs of the miR-17-92 cluster (Slaby, Svoboda et al. 2009), including miR-17-5p, which can promote β -catenin signalling (Amirkhah, Schmitz et al. 2015).
- MiR-144-3p can repress the oncogene *mammalian target of rapamycin (mTOR)* (Amirkhah, Schmitz et al. 2015).
- MiR-25-3p is potentially tumour-suppressive via the repression of *SMAD7* - which in contrast to *SMAD2* and *SMAD4* - is oncogenic, due to inhibiting TGF β 's tumour-suppressive activities (Amirkhah, Schmitz et al. 2015).
- MiR-103-3p can repress the TSG *Kruppel-like factor 4 (KLF4)* (Amirkhah, Schmitz et al. 2015).
- MiR-124-3p was shown to promote apoptosis in DLD-1 and WiDr colon cancer cells (Taniguchi, Sugito et al. 2015).

1.6.4.3 MicroRNA Interactions with the Warburg Effect

MiR-26a-5p, often overexpressed in CRC, can repress pyruvate dehydrogenase protein X component (PDHX) - a key component of the pyruvate dehydrogenase complex that catalyses the synthesis of acetyl-coenzyme A from pyruvate within the mitochondrion (Chen, Liu et al. 2014).

By inhibiting acetyl-coenzyme A synthesis, miR-26a-5p can promote a shift in metabolism such that the proportion of ATP generated via glycolysis vs. that generated by oxidative phosphorylation is increased, thereby upregulating the generation of metabolites required for mitosis (Chen, Liu et al. 2014). The promotion of the Warburg effect by miR-26a-5p has been demonstrated *in vitro* by the overexpression of the miRNA in CRC cell lines, followed by pyruvate, lactate and acetyl-coenzyme A assays.

1.6.5 Methods to Detect Differences in MicroRNA Expression between Samples

Three methods are widely employed for wide-scale miRNA comparative expression profiling:

- MiRNA array chip hybridisation.
- MiRNA array quantitative polymerase chain reaction (PCR).
- MiRNA cloning and deep sequencing.

1.6.5.1 MicroRNA Array Chip Hybridisation

MiRNA array chips are media coated with probes of sequence complementary to those of known miRNAs, to which RNA samples can be hybridised. The expression of a specific miRNA in the sample is thus evaluated according to the intensity of signal generated from its corresponding probe, which is proportional to the amount of miRNA bound. An example of such a system is the miRCURY LNA™ microRNA Arrays kit (Exiqon).

1.6.5.2 MicroRNA Array RT-qPCR

In a reverse transcription—quantitative polymerase chain reaction (RT-qPCR) ‘array’, hundreds of RT-PCRs are run in parallel, each specific for a different miRNA. An example of a system facilitating this is the TaqMan Low Density Array kit with the human miRNA Megaplex RT Primer Pool v3.0 (Slaby, Sachlova et al. 2013).

1.6.5.3 MicroRNA Library Cloning and Deep Sequencing

The small RNAs from RNA samples are captured using oligomeric adapters, then reverse transcribed to complementary DNA (cDNA) form. Specific types

of primer are used for the reverse transcription and subsequent PCR amplification of the captured small RNAs, and then the PCR products can be size-fractionated to isolate the miRNA clones, which are then subject to deep sequencing. Reads from the sequencing data are annotated based on known miRNA sequences (as catalogued in the current version of miRBase), then the relative expression of a given miRNA is evaluated by comparing its normalised read numbers between libraries. Unlike the above-mentioned miRNA array chip hybridisation and RT-qPCR methods, this approach can detect and evaluate the expression of novel miRNAs.

1.6.5.4 Comparison of Methods

The miRNA array chip hybridisation and quantitative PCR (qPCR) approaches tend to be less expensive and time-consuming than library cloning and deep sequencing methods. However, the miRNAs for which probes are included in the kits are inevitably limited to those known to exist at the time of kit development.

By contrast, library cloning and deep sequencing are able to capture novel miRNAs that may be identified as such by computational sequence analysis, and data can be re-annotated in the future against newer versions of miRBase. The list of known human miRNAs continues to grow. MiRBase v.14, which was released at the beginning of 2010, catalogued 894 human miRNAs; v.20, released halfway through 2013, 2555; and v.21, 2588. The increase in the number of mature human miRNAs catalogued by each new version of miRBase is illustrated in Figure 13 (miRBase 2014). The current version of miRBase is v. 21 and catalogues 2588 mature human miRNAs.

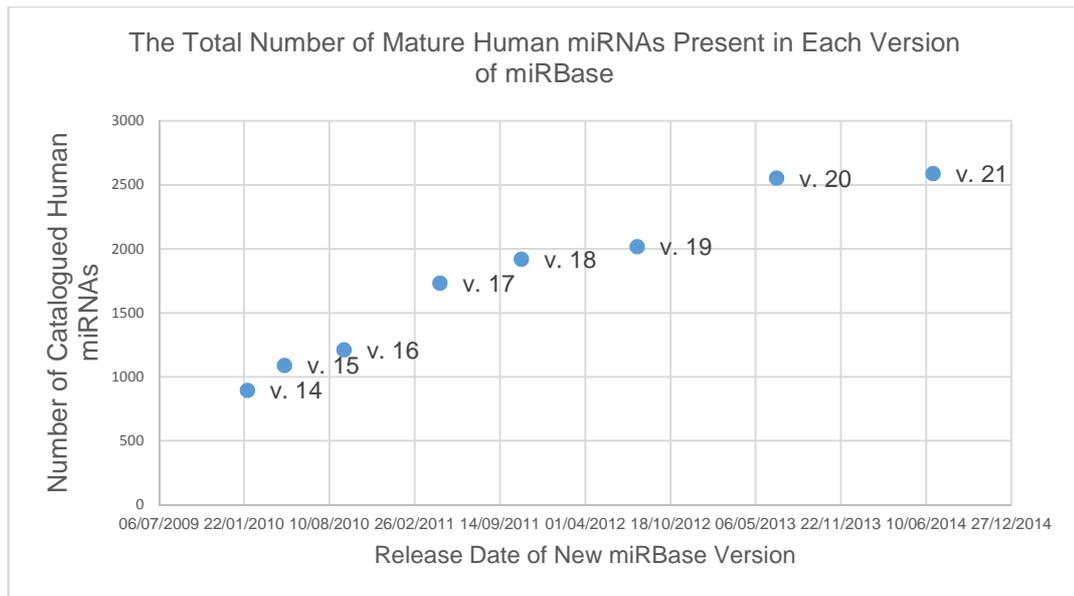


Figure 13: A graph illustrating the increase in the number of mature human miRNAs listed in miRBase with each new version, according to the miRBase Blog.

The benefits of the cloning-sequencing approach, including its greater sensitivity, potential to capture novel miRNAs, and ability to quantify miRNA expression over a wider dynamic range, were considered to outweigh its potential drawbacks of higher cost and time consumption compared with the array-based approaches. Therefore, miRNAs were to be cloned in the form of DNA-based libraries using the protocol developed by Xu et al. (Xu, Billmeier et al. 2015) using Dalmay HD adapters, and then the constructed libraries were to be subject to deep sequencing by the Earlham Institute (formerly The Genome Analysis Centre) (Norwich, UK).

1.6.5.5 Challenges in MicroRNA Cloning: Adapter Ligation Bias

The first step in miRNA cloning is to ligate adapter oligomers to the 5' and 3' ends of small RNA species. Any given adapter oligomer will inevitably vary in its affinity for different miRNAs, so those to which its affinity is greater are more likely to become ligated to the adapters. This ligation bias means if two given miRNAs are of equal abundance in a sample, but one is preferred by the adapter, that which is preferred will have more cDNA clones of itself produced and thus be read more frequently upon sequencing of the sample. Therefore, the read numbers for different miRNAs relative to one another do

not necessarily represent the actual abundances of the miRNAs relative to one another.

Assuming that the same adapter(s) is/are used for all samples, any ligation biases should occur equally across all samples and thus should not confound the evaluation of fold-differential miRNA expression between samples. However, some miRNAs may be so strongly 'disfavoured' by the ligation process that they are ultimately represented by very low or no read numbers in the sequencing data, and therefore go undetected. In this regard, ligation bias can impair sensitivity in that there may be differentially expressed miRNAs that are undetected because they are too strongly 'disfavoured' by adapter ligation.

For any given adapter sequence, ligation bias is inevitable. In order to mitigate this issue, HD adapter pools were developed by Dalmay and Sorefan et al. (Sorefan, Pais et al. 2012), which were generated by taking Illumina adapters, then adding 4 nucleotides to the miRNA-ligating ends at random. Therefore, from each single adapter are generated a potential pool of $4^4 = 256$ variants. Inevitably, each of these variants will have their own bias, but any collective bias of the pool will be much lower than that of the original single adapter. The HD adapters have been demonstrated to dramatically increase read coverage, and enable the detection of previously unidentified miRNAs (Sorefan, Pais et al. 2012).

1.7 Aims and Objectives

- To investigate the effects of SFN on miRNA expression in cancerous and normal human colonic cell lines.
 - The microenvironment of cell culture *in vitro* is not fully representative of that of tumour cells *in vivo* and cannot factor in organism-wide metabolism as can *in vivo* animal models. However, human cells cultured *in vitro* may in some instances be more representative of human tissue than 'equivalent' tissues in laboratory animals such as mice. Animal models are also more time-consuming and expensive to work with. *In vitro*

studies may involve either primary human cells or established human cell lines. Since cell lines are widely available, and experimental findings in cell lines are more likely to be more reproducible and comparable across studies, they were more appropriate than primary human cells for these studies.

- Although an array-based miRNA expression analysis by Slaby et al. (2013), using TaqMan Low Density Array qPCR, had already indicated the ability of 48 h SFN treatment to modulate the expression of several miRNAs in non-cancerous colonic cell lines (Slaby, Sachlova et al. 2013), the reported miRNA modulations had not been followed up by individual assays for confirmation. Differences in expression detected by the wide-scale analysis in this study were to be validated by single-target assays such as RT-qPCR or Northern Blots, in order to confirm that such apparent differences in expression were actually true as opposed to artefacts of the wide-scale profiling process.
- The study by Slaby et al. investigated only non-cancerous colonic cell lines, but there is a dearth of studies investigating the SFN-mediated modulation of miRNAs in their cancerous counterparts. It is conceivable that the effects of SFN in cancerous colonic cell lines differs from those in their non-cancerous counterparts, so this study therefore used both cancerous and non-cancerous colonic cell lines (Caco-2 and CCD-841 respectively).
- Most wide-scale analyses of differential expression have used array-based approaches, such as the RT-qPCR array approach adopted by Slaby et al. (2013). While relatively inexpensive and quick, this approach is limited to studying the miRNAs that were catalogued at the time of array development. The TaqMan Low Density Array used by Slaby et al. included probes for 754 miRNAs, but the current version of miRBase includes 2588 human miRNAs. This study employed the miRNA library cloning and deep sequencing approach, which is more expensive and time consuming than array-based

methods, but has the potential to detect a wider range of differentially-expressed miRNAs, including novel miRNAs.

- Challenges of the miRNA library construction process include adapter ligation bias, which can impair the sensitivity with which miRNA modulations are detected. Dalmay HD adapters were to be used in order to reduce such adapter ligation bias (Sorefan, Pais et al. 2012).
- The temporal dynamics of SFN-mediated miRNA modulation can vary widely, due to the plethora of mechanisms by which different miRNAs are able to be modulated. The steady-state level of a given miRNA depends upon the balance between its biogenesis and degradation. A specific miRNA's rate of biogenesis can be modulated at various stages, such as the transcription of the genetic loci to pri-miRNAs, the processing of pri-miRNAs to pre-miRNAs by Drosha and DGRC8, the cleavage of pre-miRNA loops by Dicer, and the loading of mature miRNAs into RISCs. Likewise, the rate at which a specific miRNA is degraded can be influenced by the degree of complementarity between the miRNA and its target mRNA, the abundance of the target, and the presence and nature of miRNA base modifications (Gantier, McCoy et al. 2011). Due to the complexities of miRNA synthesis and degradation, the sensitivity of SFN-mediated miRNA modulation detection should correlate positively with the number of different time points at which miRNA expression is profiled. However, it is impractical in terms of time and resources to profile miRNA expression at many different time points. Nevertheless, profiling expression at two different time points is feasible and expected to be significantly more sensitive than that at a single time point. This study therefore set out to profile miRNA expression at two different time points post-treatment – 8 h and 24 h.
- To investigate the potential downstream consequences of SFN-mediated changes to miRNA expression in terms of tumour suppression.

- Select miRNAs whose differential expression was validated by single-target assays were to be studied *in vitro* by the use of specific miRNA mimics and inhibitors for their ability to bind to the 3'-UTRs of predicted and/or reported mRNA targets relevant to cancer. Following the verification of such, the downstream miRNA effects on protein abundance and *in vitro* cancer 'endpoints' such as cell migration and invasion, could be investigated.

Hypothesis: SFN modulates the expression of miRNAs in both the cancerous Caco-2 and normal CCD-841 colonic cell lines, with some differential effects between the two cell lines. MiRNAs whose expression is repressed or promoted by SFN are respectively able to promote or inhibit cancer development, by directly targeting the mRNA transcripts of TSGs or oncogenes, respectively, to repress their translation.

The investigation of the above-described hypothesis by this study could help to inform the future development of anti-CRC therapeutic and/or chemopreventative strategies.

Chapter 2: Methods

2.1 Cell Culture

HCT116, Caco-2 and CCD-841 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing D-glucose (4.5g/L) and non-essential amino acids (Sigma), which was supplemented with heat-inactivated foetal bovine serum (FBS) to 10% (v/v), Penicillin-Streptomycin (Sigma) to 1% (v/v), and L-glutamine to 200nM, in 75cm² flasks. Cells were maintained in a HERAcCell® 150 CO₂ incubator at 37°C in 5% CO₂ and cell work was carried out under sterile conditions inside a Class II Biological Safety Cabinet.

HCT116 and Caco-2 cells were subcultured upon reaching 70-80% confluence. To subculture, medium was removed, cells were washed twice with 10mL pre-warmed phosphate-buffered saline (PBS), covered in 1mL trypsin-EDTA (0.25% (w/v) trypsin and 1mM EDTA), then put into 37°C incubation until all cells were detached (typically 1 min for HCT116 and 5 – 10 min for Caco-2). Then trypsin was inactivated by adding 8mL culture medium. Cell suspensions were repeatedly pipetted to disrupt cell aggregates, then cells were counted using a haemocytometer. An appropriate volume of cell suspension was then combined with fresh pre-warmed culture medium for seeding into new flasks for maintenance, and/or onto other solid media for experiments. Typically, 1.0×10^6 cells were seeded into a new 75cm² flask for maintenance, in 12mL medium.

CCD-841 cells were subcultured every 7 – 10 days in a fashion similar to that described above for Caco-2 and HCT116 cells except for a few differences. Firstly, cell monolayers were washed using Dulbecco's low-salt PBS instead of the typical PBS, prior to the addition of trypsin-EDTA. Secondly, cells were pelleted by centrifugation at 125 g for 5 min following detachment and then resuspended in fresh pre-warmed culture medium, in order to remove traces of trypsin-EDTA.

2.2 Western Blotting to Assay Proteins

2.2.1 Cell Treatments to Assay the Effects of SFN on Proteins

Cells were grown to 70-80% confluence in 6-well plates (surface area per well = 9.5cm²) and then treated by adding to culture medium dimethyl sulfoxide (DMSO)-diluted SFN (5, 10 or 20µM) or DMSO (control) for 24 h, or DMSO-diluted SFN (10µM) for 4, 8, 12, 24, 48 or 72 h; final DMSO concentrations were 0.05% (v/v).

2.2.2 Total Protein Isolation

Culture medium was removed and then cells were washed twice in ice-cold PBS, then covered in a lysis solution containing 20mM Tris-HCl (pH 8.0), 150mM NaCl, 2mM EDTA, 10% (v/v) glycerol, 1% (v/v) Nonidet P-40, cOmplete™, Mini Protease Inhibitor (Roche), and 1mM PMSF Protease Inhibitor (Thermo Fisher). Cells were incubated under the lysis solution under agitation at 4°C for 30 min, then lysates were collected by scraping wells with a rubber policeman. Collected lysates were centrifuged at 12 000 g for 15 min at 4°C and the pelleted debris was discarded.

2.2.3 Nuclear Protein Isolation

Nuclear protein fractions were isolated using the Nuclear Extract Kit (Active Motif, UK) according to the manufacturer's instructions. Briefly, culture medium was removed and then cells were washed twice in ice-cold PBS supplemented with Phosphatase Inhibitor. Cells were then covered in fresh Phosphatase Inhibitor-supplemented PBS and lifted into suspension with a rubber policeman. Suspensions were collected in pre-chilled tubes, and then centrifuged at 200 g for 5 min at 4°C to pellet the cells, discarding the supernatants. Pelleted cells were resuspended in 1X Hypotonic Buffer then put on ice for 15 min to swell.

Detergent was then added to swollen cell suspensions, which were subsequently vortexed and then centrifuged at 14 000 g for 30 sec at 4°C to separate the cytoplasmic and nuclear protein fractions, into supernatants and pellets respectively. Cytoplasmic protein-containing supernatants were transferred to separate tubes and stored at -80°C, then the nuclear protein-

containing pellets were resuspended in Complete Lysis Buffer, incubated at 4°C under rotation, vortexed for 30 sec, then centrifuged at 14 000 g for 10 min at 4°C. Nuclear protein-containing supernatants were then collected and stored at -80°C; pellets were discarded.

2.2.4 SDS-PAGE and Immunoblotting

Protein concentrations were determined using the Brilliant Blue G dye-binding assay of Bradford, after using known quantities of bovine serum albumin to generate standard curves, using Bradford Reagent (Sigma). From each sample, equal amounts of protein were combined with Loading Buffer (0.625M Tris, 2% (v/v) SDS, 10% (v/v) glycerol, 20mM DTT and bromophenol blue).

2.2.4.1 Fluorescent Detection Protocol

Protein-Loading Buffer mixtures were incubated at 95°C for 5 min to denature proteins, then loaded into 10% (v/v) sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels alongside Dual Color Precision Plus Protein™ Standards (Bio-Rad) molecular weight markers.

Proteins were subsequently size-fractionated by SDS-PAGE at 25-35mA, then transferred to low-background Immun-Blot polyvinylidene difluoride (PVDF) membranes (Bio-Rad) by semi-dry electroblotting using either the Trans-Blot SD Semi Dry Transfer Cell (Bio-Rad) or the Trans-Blot Turbo system (Bio-Rad). The SD system was run at 15V/gel for 1 h, whilst the Turbo system was run at 25V for 30 min.

Membranes were blocked by incubation in Pierce Protein—Free Blocking Buffer (Life Technologies) at room temperature for 1 h under agitation. Blocked membranes were incubated overnight under agitation at 4°C in a 1:1 mixture of the blocking buffer and PBS supplemented with 0.5% (v/v) Tween 20 detergent (PBS-T (0.5%)), in which were diluted primary antibodies against proteins of interest. Subsequently, membranes were washed for 8 min four times in PBS-T (0.1%), and then incubated for 1 h at room temperature in a 1:1 mixture of the blocking buffer and PBS-T (0.5%) supplemented with 0.002% (v/v) SDS, in which were diluted secondary

antibodies against the primary antibodies used. Membranes were subsequently washed again for 8 min four times in PBS-T (0.1%), then once in PBS for 5 min. The fluorescence of bound secondary antibody was detected using the Odyssey Imager (LI-COR), and signals were densitometrically analysed to compare protein expression between samples, using β -actin as an internal control for normalisation.

2.2.4.2 Chemiluminescent Detection Protocol

Protein-Loading Buffer mixtures were incubated at 70°C for 10 min to denature proteins, then loaded into 10% (v/v) SDS-PAGE gels alongside Dual Color Precision Plus Protein™ Standards (Bio-Rad) molecular weight markers. Proteins were subsequently size-fractionated by SDS-PAGE at 25-35mA, and then transferred to Immun-Blot PVDF membranes (Bio-Rad) by semi-dry electroblotting, using the Trans-Blot SD Semi Dry Transfer Cell (Bio-Rad) at 15V/gel for 1 h.

Membranes were blocked by incubation in 5% (v/v) skimmed milk dissolved in PBS-T (0.1%), at room temperature for 1 h 30 min under agitation. Blocked membranes were incubated overnight at 4°C under agitation, in a solution of 5% (v/v) milk, and primary antibodies against the proteins of interest, diluted in PBS-T (0.1%). Subsequently, membranes were washed for 10 min four times in PBS-T (0.1%), then incubated at room temperature for 1 h in a solution of 5% (v/v) milk, and secondary antibodies against the primary antibodies used, diluted in PBS-T (0.1%). Membranes were subsequently washed again for 10 min four times in PBS-T (0.1%).

Membranes were then developed using the ECL Western Blotting Detection Kit (GE Healthcare) according to manufacturer's instructions, by first equilibrating reagents to room temperature and then thoroughly mixing Solution A and Solution B in a 1:1 ratio. The membrane was covered in this mixture and incubated at room temperature for 5 min, whilst shielded from light. The chemiluminescence of bound secondary antibody was subsequently detected using the Fujifilm LAS3000 Imager, and signals were densitometrically analysed to compare protein expression between samples, using β -actin as an internal control for normalisation.

2.3 MicroRNA Library Construction and Sequencing

2.3.1 Cell Treatments

Cells were grown to 70-80% confluence and then treated by adding to culture medium in triplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v).

2.3.2 Total RNA Extraction

Total RNA was isolated by guanidium thiocyanate-phenol-chloroform extraction, using TRI Reagent (Sigma-Aldrich), largely according to the manufacturer's protocol with minor modifications. Briefly, culture medium was removed and then TRI Reagent was added directly onto cells (1mL per 10cm² culture dish area) to lyse them. The procedure was either paused at this stage by storing lysates at -80°C (which were fully defrosted to resume), or continued by incubating lysates at room temperature for 5 min.

200 μ L chloroform was mixed vigorously into every 1mL TRI Reagent lysate. Mixtures were incubated at room temperature for 10 min and then centrifuged at 16 000 *g* for 15 min, at 4°C, for phase separation. The aqueous (uppermost, clear) phases were then carefully transferred to fresh tubes, avoiding the other phases. RNA was precipitated from the aqueous phases by mixing in 500 μ L isopropanol and 5 μ g glycogen (co-precipitant for RNA) per 1mL TRI Reagent lysate processed, then storing at -20°C for \geq 5 h. After storage, precipitated RNA was pelleted by centrifugation at \geq 16 000 *g* for 30 min, at 4°C. Supernatants were removed from pellets, which were then washed with 1mL 75% (v/v) ethanol each. Ethanol was then removed and tubes were left open in a sterile hood to allow the evaporation of any residual ethanol (~5-10 min). Dried pellets were dissolved in water, supplemented with 10 units RNase Inhibitor (Ambion), and stored at -80°C.

2.3.3 MicroRNA Library Construction

Total RNA sample concentrations were determined using the NanoDrop spectrophotometer. Each library was constructed from 2 μ g total RNA with Dalmay's HD adapters, according to the protocol developed by Xu et al. (Xu, Billmeier et al. 2015), as described below.

2.3.3.1 Total RNA Purification

Total RNA samples were purified using the Zymo RNA Clean & Concentrator -25 kit (Cambridge Bioscience) according to manufacturer's instructions. Briefly, samples were mixed with RNA Binding Buffer and ethanol, and mixtures were centrifuged through Zymo-Spin™ IIC Columns to capture the RNA. Captured RNA was then treated with RNA Prep Buffer, washed with RNA Wash Buffer, and then eluted from the column in water. Eluted RNA was supplemented with RNase Inhibitor (Ambion) and stored at -80°C.

2.3.3.2 RNA Denaturation

2µg of RNA was combined with water to a 10µL volume, to which 1µL 3' HD Adapter was then added. The mixture was incubated at 70°C for 2 min to denature the RNA and adapter.

2.3.3.3 3' HD Adapter Ligation

Into the mixture was added 2µL T4 RNA Ligase 2 10X Buffer (NEB or Epicentre), 1µL T4 RNA Ligase 2 (NEB or Epicentre), 0.75µL RNaseOUT (40 units/µL) (Invitrogen), 4µL 50% (v/v) polyethylene glycol (PEG) (NEB), and 1.25µL water. This mixture was then incubated at 26°C for 2 h.

2.3.3.4 Degradation of Excess 3' HD Adapter

The products of the 3' HD adapter ligation reaction were purified using the Zymo RNA Clean & Concentrator -5 kit (Cambridge Bioscience) according to manufacturer's instructions, similarly to how total RNA samples were purified with the -25 kit prior to library construction, but eluting products in 13µL water.

To the purified products were added 1.6µL 10X Deadenylase Buffer (Epicentre), 1µL 5' Deadenylase (10 units/µL) (Epicentre), 0.8µL 100mM DTT, and 0.5µL RNase-OUT (40 units/µL) (Invitrogen). This mixture was incubated at 30°C for 30 min. 4µL 25mM EDTA was then added to stop the reaction.

Then, 2µL 500mM Tris-HCl (pH 9.0) (Sigma), 7µL 50mM MgCl₂, and 1µL RecJ Exonuclease (10 units/µL) (Epicentre) were added to the products. This mixture was incubated at 37°C for 30 min.

2.3.3.5 5' HD Adapter Ligation

2µL 5' HD adapter was denatured by incubating at 70°C for 2 min. Then it was mixed with the products of the excess 3' adapter degradation process, 1µL 10X T4 RNA Ligase Buffer (Epicentre), 1µL T4 RNA Ligase (Epicentre), 1µL 10mM ATP (Epicentre) and 7µL 50% (v/v) polyethylene glycol (PEG) (NEB). This mixture was incubated at 26°C for 2 h.

After the reaction was finished, 7µL water was added to the products, which were then purified using the Zymo RNA Clean & Concentrator -5 kit (Cambridge Bioscience) according to manufacturer's instructions. This was done in the same manner as the products of 3' adapter ligation were earlier purified, but eluting in 30µL water.

2.3.3.6 cDNA Synthesis

To the purified ligation products were added 4µL 10X MMLV Reverse Transcription Buffer (Epicentre), 1µL High Performance MMLV Reverse Transcriptase (Epicentre), 2µL 10mM dNTP PreMix, 2µL 100mM DTT (Epicentre), and 1µL Reverse Transcription Primer (20µM) (Sigma). The mixture was incubated at 37°C for 20 min, and then the reaction was terminated by incubating at 85°C for 15 min.

2.3.3.7 PCR Amplification

4µL reverse transcription product was mixed with 0.5µL 10mM dNTPs, 4µL 5X High Fidelity Phusion Buffer (Thermo Scientific), 0.2µL Phusion DNA Polymerase (2 units/µL) (Thermo Scientific), 1µL Illumina RP-1 Primer (10µM) (Sigma), 1µL Illumina Index Primer (10µM) (Sigma), and 9.3µL water.

This mixture was then incubated in a thermocycler with the following programme:

1. 98°C for 30 sec
2. x cycles of [98°C for 10 sec, 60°C for 30 sec, 72°C for 15 sec], where $12 \leq x \leq 16$.
3. 72°C for 10 min

Different numbers of cycles (x) were tried at once for each library, in order to determine which number was optimal.

2.3.3.8 PAGE Separation and Elution

The PCR products were size-fractionated by PAGE alongside DNA size markers to separate the miRNA-representing polynucleotides from those representing other RNA species of differing size. Electrophoresis was run in an 8% (v/v) gel at 120V for 2 h, which was then stained with SYBRGold, and scanned to visualise the location of the miRNA bands i.e. those located between 120 and 140 bp (since their sequences were products of 3' and 5' HD adapters, miRNAs and PCR primers). Pieces of gel corresponding to these bands were physically cut out.

Gel-breakers were made by piercing the bottoms of 0.5mL Eppendorf tubes with a 21-gauge needle, then putting them inside 1.5mL Eppendorf tubes. Cut out pieces of gel were transferred to the 0.5mL compartments of the gel-breakers, which were then centrifuged to force the gel through the holes, thus breaking it into small fragments. Then 200-400 μ L 1X Elution Buffer (NEB) was mixed into the gel fragments, and mixtures were incubated either at room temperature for 2 h, or 4 $^{\circ}$ C overnight, with shaking.

Gel fragment-buffer mixtures were then transferred to Spin-X filtration columns, which were centrifuged for 5 min to separate dissolved DNA from the gel debris. DNA was recovered from solutions by ethanol precipitation, by adding 2 μ L GlycoBlue, 30 μ L 3M NaOAc, and 975 μ L ethanol, to each 300 μ L solution. Mixtures were then stored at -80 $^{\circ}$ C for \geq 2 h, then centrifuged at 16 000 g for 30 min to pellet precipitates. Pellets were washed in 75% (v/v) ethanol, dried, and then dissolved in 12 μ L water each. The solutions constituted the final constructed libraries.

2.3.4 Deep Sequencing

For each cell line, nine miRNA libraries were mixed in a pool which was then subjected to deep sequencing by the Earlham Institute on the Illumina HiSeq 2500 platform, with 100 million reads. Relative library concentrations were compared by PAGE of 1 μ L of each library, followed by densitometric band analysis. Based on this densitometry, equal amounts of each library were pooled together.

2.3.5 Data Analysis (Performed by Claudia Paicu)

Data were converted to FASTA format then quality-checked. Sequences containing 'N's were discarded. Adapter-adapter dimer sequences were removed from the data, and the common adapter sequence (TGGAATTCT) and *NNNN* ends of the HD adapters were trimmed from remaining read sequences. The quality of the data was assessed by checking for expected size class distributions and complexity (see Appendix for graphs and further detail). Replicate consistency was evaluated according to the Jaccard index, MA plots and box plots (see Appendix for data, plots and further detail).

Non-redundant sequence data files were created and then sequences were mapped full-length against the human genome using PatMan (allowing no gaps or mismatches) after eliminating low-complexity reads. Any libraries in which a high proportion of reads did not match the human genome, and/or inter-replicate consistency was low, were disregarded. Remaining library data were initially normalised by three different methods: bootstrapping, reads-per-million normalisation and quantile normalisation. Whichever method implied the greatest inter-replicate consistency was considered to be the most appropriate to use.

Read sequences were then annotated against the mature human miRNA sequences listed in MiRBase using PatMan, permitting ≤ 2 mismatches (0 gaps) to account for post-transcriptional miRNA modifications and/or sequencing errors. MiRNA read numbers were then normalised by the method previously determined as most appropriate for the data.

Most miRNAs were represented by multiple sequences, differing mostly by length. All reads of sequences annotated as the same miRNA were summed together. Fold changes in miRNA expression were estimated based upon the median values of replicates, with an offset of 20 added to each to reduce noise from very-low-abundance sequences. For example, if (0 h 1 = 11; 0 h 2 = 13; 0 h 3 = 9; 8 h 1 = 34; 8 h 2 = 33; 8 h 3 = 35), fold change from 0 to 8 h = $(\text{Median}(8\text{ h})+20)/(\text{Median}(0\text{ h})+20) = (34+20)/(11+20) = 1.7$. Read numbers for each miRNA were compared between 0 h and 8 h, 0 h and 24 h,

and 8 h and 24 h; miRNAs with a reported fold change ≥ 1.9 in at least one of the comparisons were regarded as potential SFN-modulated miRNAs.

2.4 Northern Blotting to Validate the Effects of SFN on MicroRNA Expression

2.4.1 Cell Treatments

Cells were grown to 70-80% confluence in 10cm cell culture dishes (surface area = 55cm²), and then treated in duplicate by adding to culture medium, DMSO alone (control) or DMSO-diluted SFN (10 μ M) for 8 or 24 h (final DMSO concentrations were 0.05% (v/v)).

2.4.2 Total RNA Extraction

Total RNA was isolated by guanidium thiocyanate-phenol-chloroform extraction, using TRI Reagent (Sigma-Aldrich), largely according to the manufacturer's protocol with minor modifications. Briefly, culture medium was removed and then TRI Reagent was added directly onto cells (1mL per 10cm² culture dish area) to lyse them. The procedure was either paused at this stage by freezing lysates at -80°C (defrosted to resume), or continued.

Lysates were incubated at room temperature for 5 min, then 200 μ L chloroform was mixed in vigorously per 1mL TRI Reagent. Mixtures were incubated at room temperature for 10 min, and then centrifuged at 16 000 g for 15 min at 4°C for phase separation. The aqueous (uppermost and clear) phases were carefully transferred to fresh tubes, avoiding the other phases. RNA was precipitated from the aqueous phases by adding 500 μ L isopropanol and 5 μ g glycogen per 1mL TRI Reagent lysate processed. Mixtures were stored at -20°C for ≥ 5 h, before centrifuging at ≥ 16 000 g for 30 min, at 4°C, to pellet precipitated RNA. Supernatants were removed from pellets, which were then washed in 1mL 75% (v/v) ethanol each. Ethanol was then removed and tubes were left open in a sterile hood to allow the evaporation of residual ethanol (~5-10 min). Dried pellets were dissolved in water, supplemented with 10 units RNase Inhibitor (Ambion), and then stored at -80°C.

2.4.3 Small RNA Extraction

Small RNA fractions were isolated using the MirPremier™ MicroRNA Isolation Kit (Sigma-Aldrich) according to the manufacturer's instructions.

Briefly, culture medium was removed from cells, which were then washed twice with PBS, and then covered in a lysis mixture (MicroRNA Lysis Buffer and Binding Solution combined in a 7:3 ratio). Cells were left under the lysis mixture for ≥ 5 min with occasional rocking of plates, then lysates were collected and centrifuged at 16 000 g for 5 min to pellet cell debris, genomic DNA and larger RNA species. Small RNA-containing supernatants were collected, mixed with 1.1 volumes of ethanol, and then centrifuged through a Binding Column to capture RNA. Captured RNA was washed with ethanol then with Wash Solution 2, and finally eluted in 50 μ L Elution Solution.

RNA was then concentrated by ethanol precipitation, which was done by adding 2 vol. ethanol, 0.1 vol. NaOAc (3M, pH 5.2) and 5 μ g glycogen, then storing at -80°C for ≥ 2 h. After storage, mixtures were centrifuged at 16 000 g for 30 min, at 4°C, to pellet precipitates. Supernatants were removed and then pellets were washed in 70% (v/v) ethanol, before removing ethanol, allowing residual ethanol to evaporate in a sterile hood, and then dissolving dried pellets in water. RNA solutions were supplemented with 10 units RNase Inhibitor (Ambion) and then stored at -80°C.

2.4.4 RNA Electrophoresis and Blotting

PAGE equipment was thoroughly cleaned using detergent to minimise the presence of RNases, then 16% (v/v) Urea-PAGE gels were made. RNA concentrations were determined using the NanoDrop spectrophotometer, then equal amounts of RNA were combined with 2X Gel Loading Buffer II (Ambion). Mixtures were heat-denatured at 65°C for 5 min to denature RNA. The Urea-PAGE gel was pre-run at 80V for 20 min, and then the RNA-buffer mixtures were loaded into the gel. Electrophoresis was run at 100V for 2 h, in 0.5X tris/borate/EDTA (TBE) buffer, to size-fractionate RNAs.

Then RNA was transferred from gels to Hybond NX membranes by semi-dry electroblotting in 0.5X TBE. This was done at a constant current of 3mA per

cm² transfer area for 1 h, in a 'sandwich' of Whatmann filter papers. After RNA transfer, the procedure was either paused by storing membranes at -20°C, or continued.

RNA was chemically crosslinked to membranes using carbodiimide. The crosslinking solution was prepared by adding 122.5µL of 12.5M 1-methylimidazole and 10µL of 1M HCl to 10mL water, and then dissolving 0.373g 1-ethyl-3-(dimethylaminopropyl) carbodiimide (EDC) powder into it. The crosslinking solution was used to soak a Whatmann filter paper, onto which the membrane was placed (with the RNA side facing away from the paper). The membrane and filter paper were then sealed in clingfilm and incubated at 60°C for 1-2 h. After incubation, membranes were washed in water for 10 min, then stored at -20°C.

2.4.5 Probe Hybridisation and Signal Detection

PerfectHyb hybridisation buffer was pre-incubated at 37°C to dissolve any precipitate. Membranes were incubated at 37°C in 10mL hybridisation buffer each, under rotation, for 1-2 h. Radioactively-labelled probes were prepared from DNA oligonucleotides of sequence anti-sense to those of miRNAs of interest, by mixing 4µL 5X Forward Reaction Buffer (Thermo Scientific), 2µL DNA oligonucleotide (10µM), 1µL T4 kinase enzyme (Thermo Scientific), 10-11µL water and 2-3µL γ-ATP (P-32), then incubating at 37°C for 1 h.

Prepared probes were added directly into the hybridisation buffer in which the membranes were incubated. Membranes were incubated in the probe-containing hybridisation buffer at 37°C overnight under rotation. Afterwards, the probe-containing buffer was discarded, and membranes were washed in 0.2X SSC/0.1% (v/v) SDS solution. Firstly, they were shaken vigorously in the solution for 10 sec; secondly they were incubated with fresh solution at 37°C under rotation for 15 min; thirdly, the solution was discarded and the rotating incubation was repeated with fresh solution; and finally, they were shaken vigorously in solution for 10 sec again. After washing, membranes were sealed in clingfilm and then placed onto phosphorimaging plates (RNA sides of membranes facing the reactive sides of plates), inside cassettes, for overnight exposure at 4°C.

Exposed phosphorimaging plates were scanned to detect signals generated by radiation-induced excitations of the plate, and the signal data were densitometrically analysed using the ImageJ software to compare relative miRNA expression between samples, using U6 as an internal loading control. Before re-probing, membranes were stripped of bound probe either by incubating in 0.1% (v/v) SDS at 95°C under rotation until radioactive signal could no longer be detected using the Geiger counter (usually 1-2 h), or by placing membranes into boiled 0.1% (v/v) SDS until cooled to room temperature (~45 min). Stripped membranes were returned to storage at -20°C.

2.4.6 Dot Blots to Test Probe Cross-Reactivity

In order to screen for potential cross-reactivity between probes and off-target miRNAs similar in sequence to the target, dot blots were prepared by pipetting serial dilutions of positive sense DNA oligomers of sequence equivalent to those of different miRNAs onto a Nylon NX membrane, and then crosslinking the DNA to the membrane in the same manner that RNA was crosslinked to membranes for Northern Blot preparation. Dot blots were then hybridised using anti-sense oligo probes as per the Northern Blot protocol. The amounts of positive sense oligo used ranged between 1fmol and 1pmol, in 1µL volumes.

2.5 TaqMan® RT-qPCR to Assay the Effects of SFN on Messenger RNA

2.5.1 Cell Treatments

Cells were grown to 70-80% confluence in 6-well plates (surface area/well = 9.5cm²), then treated by adding to culture medium in triplicate, DMSO (control) or DMSO-diluted SFN (5, 10 or 20µM) for 24 h, or DMSO-diluted SFN (10µM) for 4, 8, 12, 24 or 48 h; final DMSO concentrations were 0.05% (v/v).

2.5.2 RNA Extraction

RNA was extracted using the GenElute™ Total Mammalian RNA Miniprep Kit (Sigma, UK) according to manufacturer's instructions. Briefly, culture medium was removed from cells, which were then covered in Lysis Buffer. Lysates were centrifuged through a Filtration Column, and then the flow-through liquid was mixed with ethanol and then centrifuged through an RNA binding column to capture the RNA on the column. Captured RNA was then washed with Wash Solution 1 then Wash Solution 2, eluted in Elution Solution, supplemented with 10 units of RNase Inhibitor (Ambion), and then stored at -80°C.

2.5.3 Reverse Transcription

RNA concentrations were determined using the NanoDrop spectrophotometer (Thermo Scientific). Equal amounts of RNA were reverse transcribed to cDNA using the qScript SuperMix (Quanta BioSciences, USA) according to manufacturer's instructions, heating reactions in a thermocycler at 25°C for 5 min, 42°C for 30 min, then 85°C for 5 min.

2.5.4 Primers and Probes for TaqMan® RT-qPCR

TaqMan® primers and probes had been designed previously by the Bao laboratory group with Primer Express based on the sequences of the human mRNAs as described in Table 3.

Table 3: The sequences of the primers and probes used for RT-qPCR.

Target	Component	Sequence	Conc. (nM)
18S	Forward Primer	5'-GGCTCATTAATCAGTTATGGTTCCT-3'	200
	Reverse Primer	5'-GTATTAGCTCTAGAATTACCACAGTTATCCA-3'	200
	Probe	5'-TGGTCGCTCGCTCCTCTCCCAC-3'	100
TrxR-1	Forward Primer	5'-CCACTGGTGAAGACCACGTT-3'	200
	Reverse Primer	5'-AGGAGAAAAGATCATCACTGCTGAT-3'	300
	Probe	5'-CAGTATTCTTTGTCACCAGGGATGCCCA-3'	100

Probes were labelled with fluorescent reporter 6-carboxyfluorescein (FAM) and the tetramethylrhodamine (TAMRA) quencher at their 5' and 3' ends respectively.

2.5.5 RT-qPCR

PCR amplification was carried out in 96-well PCR plates using the TaqMan® Universal PCR Master Mix (Applied Biosystems), according to manufacturer's instructions, with the following PCR programme:

- 95°C for 10 min
- 45 cycles of:
 - 95°C for 15 sec
 - 60°C for 1 min

96-well PCR plates were used, and generated fluorescent signals were detected in real-time using the ABI 7500 Fast Real-Time PCR System (Applied Biosystems). Standard curves and 'no template controls' were included for each target.

2.5.6 Data Analysis

Relative mRNA expression was evaluated based on C_T values (i.e. the number of cycles required before a defined threshold is reached) by the Pfaffl method, which takes into account the replication efficiency of each target as indicated by the standard curve data. 18S was used as an internal control for normalisation.

Pfaffl's Equation:

Where A = amplification factor = efficiency + 1; G = transcript of interest; 18S = internal reference;

$$\text{Relative Expression}(G[\text{Treated}]) = \frac{A[G]^{(CT[G(\text{Control})]-CT[G(\text{Treated})])}}{A[18S]^{(CT[18S(\text{Control})]-CT[18S(\text{Treated})])}}$$

2.6 MicroRNA Assays by TaqMan® RT-qPCR

2.6.1 Cell Treatments

Cells were grown to 70-80% confluence in 6-well plates (surface area per well = 9.5cm²) then treated by adding to culture medium in triplicate, DMSO (control) or DMSO-diluted SFN (5, 10 or 20µM) for 24 h, or DMSO-diluted SFN (10µM) for 4, 8, 12, 24 or 48 h (final DMSO concentrations were 0.05% (v/v)).

2.6.2 Small RNA Extraction

Small RNA was extracted using the column-based miRPremier™ MicroRNA Isolation Kit (Sigma-Aldrich), as described above (see Northern Blotting to Validate the Effects of SFN on MicroRNA Expression: Small RNA Extraction).

2.6.3 Total RNA Extraction

Total RNA was extracted by guanidium thiocyanate-phenol-chloroform extraction using TRI Reagent (Sigma-Aldrich) as described above (see MicroRNA Library Construction and Sequencing: Total RNA Extraction).

2.6.4 Reverse Transcription

First strand cDNA for specific targets, including the internal control U6, was synthesised using the specific TaqMan® Small RNA Assay kit according to the manufacturer's instructions, and TaqMan® Small RNA Assay kit primers specific to each target (Applied Biosystems). Reactions were heated in a thermocycler at:

- 16°C for 30 min
- 42°C for 30 min
- 85°C for 5 min

Relative mRNA abundance compared with controls was evaluated by the Pfaffl method, taking into account different efficiencies of replication during PCR as determined from standard curve data, using U6 as the internal control.

Pfaffl's equation:

Where A = amplification factor = efficiency + 1; G = transcript of interest; U6 = reference transcript;

$$Relative\ Expression(G[Treated]) = \frac{A[G]^{(CT[G(Control)]-CT[G(Treated)])}}{A[U6]^{(CT[U6(Control)]-CT[U6(Treated)])}}$$

2.7 Cloning 3'-UTRs into Vectors for Luciferase Assays

2.7.1 Making the Wild-Type 3' UTR Clones

The 3'-UTR sequences of genes of interest were obtained from Ensembl, and parts of 3'-UTRs containing the predicted miRNA binding sites were amplified from human genomic DNA by PCR with the In Fusion HD Cloning Plus Kit (Clontech) according to manufacturer's instructions, using custom designed PCR primers.

Each designed PCR primer consisted of an invariant manufacturer-specified vector-specific sequence (VSS) at the 5' end – containing the SacI restriction site – and a custom gene-specific sequence (GSS) at the 3' end. GSSs were designed using Primer-BLAST, aiming for a length of 18-25 bases, a GC content of 40-60%, and a melting temperature (T_M) of 58-65°C. Differences in T_M between the GSSs of forward and reverse primers in each pair were $\leq 5^\circ\text{C}$. T_M s were calculated with the Promega BioMath Calculator. Oligocalc was used to check that full primers would not exhibit high self-complementarity.

The manufacturer-specified VSSs to allow ligation at the SacI site of the vector (5' to 3') were:

Forward	CTAGTTGTTTAAACG
Reverse	GACTCGAGGCTAGCG

The sequences of the designed PCR primers are described in

Table 4.

Table 4: The sequences of the designed PCR primers.

3' UTR Target	FW/RV	Designed Primer Sequence (VSS Underlined)	Product Length (bp)
BACH1	FW	<u>CTAGTTGTTTAAACGGGTGTAG</u> GGGAGGATATTGC	1239
	RV	<u>GA</u> CTCGAGGCTAGCGCACATG CAACCTGTATCGTCT	
CDC25A[a]	FW	<u>CTAGTTGTTTAAACGGTGACATT</u> TGGAGAGGGGGC	108
	RV	<u>GA</u> CTCGAGGCTAGCGACAGGG ACAGAAGAGGCGTA	
CDC25A[b]	FW	<u>CTAGTTGTTTAAACGCAGGGCA</u> TCTTGCTGGCTAC	1613
	RV	<u>GA</u> CTCGAGGCTAGCGGGCAGA GAGCATGGGTTCAA	
HMGA2[a]	FW	<u>CTAGTTGTTTAAACGGTTCGATT</u> TCTACCTCAGCAGC	1793
	RV	<u>GA</u> CTCGAGGCTAGCGCGTGTTCC CTTTCTATCAAATGTC	
HMGA2[b]	FW	<u>CTAGTTGTTTAAACGGCTGCTT</u> CAGGGAGGTAGTT	479
	RV	<u>GA</u> CTCGAGGCTAGCGGCGACC AACAAACAGCAAAGAA	
KRAS	FW	<u>CTAGTTGTTTAAACGGGTCTCT</u> GTGCCAGCTCTA	757
	RV	<u>GA</u> CTCGAGGCTAGCGTGCCTAC TTGGGAACATTCACT	
MYC	FW	<u>CTAGTTGTTTAAACGAATGTCCT</u> GAGCAATCAC	112
	RV	<u>GA</u> CTCGAGGCTAGCGGAGGCA GTTTACATTATGGC	

Human gDNA was isolated from CCD-841 cells using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma) according to the manufacturer's instructions. PCRs were set up by mixing 12.5µL CloneAmp HiFi PCR Premix, 0.6µL FW primer, 0.6µL RV primer, 2.0µL gDNA (44ng/µL) and 9.3µL water. The following PCR programme was run: 98°C for 10 sec,

followed by 35 cycles of [98°C for 10 sec; x°C for 15 sec; 72°C for y sec], where $x = 0.3(\text{lowest } T_M \text{ of primer pair}) + 0.7(\text{Product } T_M) - 14.9$; $y \approx 5(\text{product length in kilobases})$. PCR products were visualised by electrophoresis in a 1.5% (w/v) agarose gel (60V for 1 h, with ethidium bromide), alongside a DNA marker (“100bp”, NEB). The PCR product lane was checked for the presence of a single band, indicating the specific amplification of one product, and the size of the product was checked against the marker.

Alongside the 3'-UTR inserts, an insert containing several repeats antisense to the whole sequence of let-7f-5p was created, using two synthetic complementary oligomers (118 nucleotides in length) and annealing them together. The “sense” strand oligo contained the repeats antisense to let-7f-5p and was of sequence:

```
CTAGTTGTTTAAACGAACTATAACAATCTACTACCTCAAACCTATAACAATCTA  
CTACCTCAAACCTATAACAATCTACTACCTCAAACCTATAACAATCTACTACCT  
CACGCTAGCCTCGAGTC.
```

The duplex was created by first purifying 2µL of each oligomer using the DNA Clean & Concentrator -5 kit (Zymo) according to the manufacturer's instructions, eluting in 10µL supplied buffer each. Then 8µL of each eluate (concentration $\approx 75\text{ng}/\mu\text{L}$) was combined with 4µL sodium chloride (10mM) and 20µL water in a single mixture. The mixture was incubated at 95°C for 5 min, then the incubation temperature was reduced by 0.1°C/sec until 25°C was reached. The product was purified again with the DNA Clean & Concentrator -5 kit, and the presence of a single product of correct size was verified by agarose gel electrophoresis.

The pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega) was linearised by digestion with the SmaI restriction enzyme, and then purified using the DNA Clean & Concentrator -5 kit (Zymo), according to the manufacturer's instructions. Verified 3'-UTR PCR product inserts were then treated with Cloning Enhancer (In Fusion HD Cloning Plus Kit, Clontech) by mixing 5µL PCR product with 2µL Cloning Enhancer, then incubating at 37°C for 15 min and then at 80°C for 20 min. Then 2µL of this mixture was mixed

with 2µL 5X In Fusion HD Enzyme Premix, 50-200ng linearised pmiRGLO vector, and water up to a 10µL volume. This mixture was incubated at 50°C for 15 min, and then stored at 4°C (short-term) or -20°C (long-term).

The ligation of “let-7f-5p repeat” duplex into linearised vector was performed similarly, except that the Cloning Enhancer treatment was skipped.

2.7.2 Growing the Wild-Type Vectors in *E. coli*

Competent DH5α *E. coli* cells were defrosted from -80°C storage, to be transformed with the vector ligation products. 5µL ligation product was mixed with 100µL defrosted bacterial suspension, then mixtures were put onto ice for 30 min. Then they were heat-shocked at 42°C for 45 sec in a heat-block, before further placement on ice for 2 min. 400µL lysogeny broth (LB) was then added per 100µL bacterial suspension, and mixtures were incubated at 37°C for 1 h with shaking. LB-agar plates (with ampicillin) were prepared, onto which the bacteria-LB mixtures were spread following incubation (40-100µL mixture over each plate). Spread plates were then incubated at 37°C overnight.

Following overnight incubation, 10 colonies were picked from each plate, and used to separately inoculate 5mL aliquots of LB, which were then incubated at 37°C overnight with shaking. A₆₀₀ of cultures between 2.0 and 4.0 indicated that they were at a suitable stage of growth for plasmid extraction. Then, either cultures were put into 4°C storage, or plasmids were immediately isolated using the column-based QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer’s instructions.

Briefly, bacteria were pelleted by centrifugation and then resuspended in Buffer P1 (supplemented with RNAse A). Then the suspensions were mixed with Buffer P2 to lyse the cells, and then Buffer N3 was added to stop the lysis reactions. Mixtures were centrifuged at 16 000 g for 10 min to pellet debris, then supernatants were centrifuged through QIAprep spin columns to capture DNA, which was then washed with Buffer PB, then with Buffer PE. Washed DNA was eluted from the columns in 50µL Buffer EB each and then stored at 4°C (short-term) or -20°C (long-term).

2.7.3 Testing Prepared Vectors for Presence of 3'-UTR Inserts in Correct Orientation

The ligation of 3'-UTR inserts into the pmiRGLO vector is not 100% efficient; the products are a mixture of vectors with successful insertions, and vectors ligated back together without any insert. Thus, some of the bacterial minipreps will consist of the desired vector with insert, but some of vector without insert. The presence/absence of insert was determined by digestion with restriction enzymes such as SacI, followed by agarose gel electrophoresis (1% (w/v) gel; gel and running buffer both supplemented with 6.25 μ L EtBr/100mL) against a '1kb' DNA Ladder (NEB); SacI digestion of vector with the insert is expected to excise the insert, thus producing two linear fragments: one of the insert itself and the other of the rest of the vector. By contrast, digestion of vector without insert is expected to cut the vector once, thus forming one linear product only. However, SacI digestion usually failed, perhaps due to disruption of the SacI sites during the ligation process. Therefore, vectors were instead digested with HindIII, for which there are two restriction sites in the pmiRGLO vector, flanking the SacI site at which the UTRs are inserted. In this instance, two linear products are produced from digestion: one of 4771bp and another of 2579bp if no insert is present, or of 2579+[length of insert]bp. For some short inserts such as those for CDC25A[a] and MYC, plasmids were sequentially digested with both HindIII and NheI; NheI cleaved the HindIII-generated (2579+[length of insert]bp) product to 867bp and (1710+[length of insert]bp) fragments; the position of the latter being used to determine the presence or absence of the insert (see Appendix for further detail).

The vectors verified as containing the desired inserts were then sent for Sanger sequencing (Eurofins) to check that the insertion occurred in the desired orientation. To prepare the samples for sequencing, 15 μ L volumes of each were prepared at concentrations of 100ng/ μ L, then 2 μ L of pmiRGLO forward sequencing primer (10 μ M) was added, of the following sequence:

5'-CATTAAGGCCAAGAAGGGCGGC-3'.

Additionally, samples were similarly prepared each with the pmiRGLO reverse sequencing primer (10µM) for reverse sequencing:

5'-CCTTTCGGGCTTTGTTAGCAGC-3'

2.7.4 Making Vectors with Mutated MicroRNA-Target Sites

Mutations were directed to the seed regions in the predicted miRNA-target sites, such as to change them to restriction sites, whose restriction enzymes would cut the wild-type construct once only, as reported by Restriction Mapper. Forward primers for mutagenesis PCR were designed with regards to the guidelines for optimal primer design according to the Quikchange Mutagenesis Kit Manual (Agilent Technology), aiming to have the mutated site towards the middle, a $T_M \geq 78^\circ\text{C}$ and a GC content $\geq 40\%$, where:

$$T_m = 81.5 + 0.41(\%GC) - \frac{675}{Length} - \%mismatch$$

In the above equation, %GC and %mismatch are both input as rounded integers (e.g. 43.9% as 44). Please see the Appendix for the full primer sequences. The mutations were directed as illustrated in Table 5.

Table 5: A summary of where and what mutations were directed to.

Clone	Site	Targeted Site (Wild-Type)	Targeted Site After Mutation	Newly Created Restriction Site
BACH1	i	CTACCT	GATATC	EcoRV
	ii	TACCTC	GGATCC	BamHI
CDC25Aa	i	TACCTC	GATATC	EcoRV
CDC25Ab	i	TACCTC	GGATCC	BamHI
	ii	TACCTC	CAATTG	MfeI
HMGA2a	i	TACCTC	GGATCC	BamHI
	ii	TACCTC	GAATTC	EcoRI
	iii	TACCTC	CAATTG	MfeI
	iv	TACCTC	ACGCGT	MluI
	v	TACCTC	CATATG	NdeI
HMGA2b	vi	TACCTC	GATATC	EcoRV
	vii	TACCTC	GGATCC	BamHI
	viii	TACCTC	AGATCT	BglII
KRAS	i	TACCTC	GATATC	EcoRV
MYC	i	CTGCCT	GGATCC	BamHI

Sequences for the corresponding reverse primers were derived by taking the reverse complement of the designed forward primer sequences. PCR-based

mutagenesis was carried out using 'wild-type' constructs as the templates. PCRs were set up by mixing 5µL 10X Pfu Ultra II Reaction Buffer, 1µL Pfu Ultra II HS DNA Polymerase (2.5 units/µL), 25ng 'wild-type' construct, 1µL dNTP Mix (10mM), 125ng FW primer (10µM), 125ng RV primer (10µM), and water up to a 50µL volume. The PCR programme consisted of 95°C for 30 sec, followed by 18 cycles of [95°C for 30 sec; 55°C for 1 min; 68°C for x min], and then 68°C for 10 min, where x = length of plasmid in kilobases. PCR products were purified using the Zymo DNA Clean & Concentrator kit (Zymo Research) according to manufacturer's instructions.

Remaining 'wild-type' vector templates in the purified products were destroyed using the DpnI restriction enzyme, which digests into fragments only methylated DNA (the 'wild-type' vector from bacteria is methylated, but the mutagenesis PCR products are unmethylated). Then competent DH5α *E. coli* were transformed with the products, as done previously for the wild-type constructs (see above). Vector-containing bacteria were selected with ampicillin as before, and minipreps were similarly isolated. Isolated minipreps were checked for successful mutation of the binding site(s) by digestion with a restriction enzyme targeting the new site that successful mutation would generate, and by which the non-mutated vector would be cut once only. Digestion products were then analysed by agarose gel electrophoresis (1% (w/v) gel; gel and running buffer both supplemented with 6.25µL EtBr/100mL) (see Appendix for images). Those indicated to be successful mutants according to digestion were then subject to Sanger sequencing (Eurofins) for verification (see the Appendix for sequencing data).

One predicted binding site per construct was mutated at once, beginning with the first of each. A successful mutant was used as the template for mutating the next seed region.

2.8 Transfection and Luciferase Reporter Assays

2.8.1 Transfection

HCT116 cells were simultaneously seeded in 96-well plates (surface area per well = 0.29cm²) with 3.0×10^4 cells/well, and 'fast-forward' transfected in

triplicate with luciferase reporter vector constructs along with miScript miRNA Mimic (Qiagen) and/or miScript miRNA Inhibitors (Qiagen) and/or AllStars Negative Control siRNA (Qiagen) using the Attractene Transfection Reagent (Qiagen), according to the manufacturer's instructions.

Briefly, cells were detached as for routine subculture, resuspended in complete medium, and then seeded into 96-well plates with volumes per well of 100 μ L. Per triplicate transfection, a 165 μ L transfection mixture was made by diluting 2.48 μ L Attractene, 660ng vector (and miRNA mimics and/or inhibitors and/or AllStars Negative Control siRNA at desired concentrations) in non-supplemented DMEM. Transfection mixtures were incubated at room temperature for 15 min, and then 50 μ L of each was mixed into each existing 100 μ L cell suspension, so that final well volumes were 150 μ L. Additionally, 3 wells of cells had 50 μ L non-supplemented DMEM added to them, for background luminescence measurements. Cells were then incubated under typical culture conditions for 48 h.

2.8.2 Luciferase Reporter Assays

The expression of the firefly luciferase is inhibited by the binding of miRNAs or their mimics to the 3'-UTR target sites that have been inserted into the pmiRGLO vector, whereas that of *Renilla* luciferase is not affected by this. Therefore, *Renilla* luminescence is a suitable internal control against which to normalise firefly luminescence, in order to account for any well-to-well variation in cell viability and/or nucleic acid transfection efficiency.

Luciferase Assay Reagent was prepared by dissolving all supplied Dual-Glo[®] Luciferase Substrate in the Dual-Glo[®] Luciferase Buffer at once, and the prepared reagent was stored at -80°C between uses. The Stop & Glo[®] Reagent was prepared immediately before use by mixing Dual-Glo[®] Stop & Glo[®] Substrate with Dual-Glo[®] Stop & Glo[®] Buffer in a 1:100 ratio.

Reagents and cells were equilibrated at room temperature before beginning the assay reactions.

After the 48 h incubation, 75 μ L medium was removed from each well and discarded, and then 75 μ L of Dual-Glo[®] Luciferase Assay Reagent (Promega) was mixed into the remaining suspension, to both lyse the cells

and to provide substrate for the firefly luciferase. After ≥ 15 min incubation at room temperature, luminescence was measured using the FLUOstar Omega microplate reader (BMG Labtech), and data were blank-corrected based on measurements from the non-transfected wells to which non-supplemented DMEM instead of transfection mixture had been added.

After these measurements, 75 μ L Stop & Glo Reagent (Promega) was mixed into each well, to both quench the firefly luminescence and to provide substrate for *Renilla* luciferase. Following further incubation at room temperature for ≥ 15 min, luminescence was measured and blank-corrected similarly to as before.

2.9 Clonogenicity Assays

2.9.1 First Method

Cells were seeded in 6-well plates (surface area/well = 9.5cm²) at low density (400 cells/well), allowed to attach for 48 h, and then treated by adding to culture medium in sextuplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Following the addition of SFN, cells were incubated for 14 days to allow colony formation, after which cells were fixed by washing in PBS and then incubating under 4% (v/v) paraformaldehyde (PFA) for 10 min. Fixed cells were washed in PBS and then stained by incubating under Crystal Violet solution (0.1% (v/v) in water) for 30 min. Stained cells were washed in water and then the number of colonies in each well was counted under a microscope.

2.9.2 Second Method

Cells were grown to 70-80% confluence in 24-well plates (surface area per well = 1.9cm²), then treated by adding to culture medium in 6 replicates, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). 24 h following the addition of SFN, cells from each well were detached using trypsin, then 400 from each were seeded into wells of a 6-well plate (surface area per well = 9.5cm²).

Subsequently, cells were incubated for 14 days to allow colony formation, after which cells were fixed by washing in PBS then incubating under 4% (v/v) PFA for 10 min, and then stained by washing fixed cells in PBS then incubating under Crystal Violet solution (0.1% (v/v) in water) for 30 min. Stained cells were washed in water and then the number of colonies in each well was counted.

2.10 Cell Viability Assays

Cell viability was measured using the MTT assay, which is based upon measuring the optical densities of DMSO-solutions of purple formazan, the solutes of which are generated by the metabolic reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) by viable cells and precipitated. Equal amounts of DMSO are then added to each well to dissolve the precipitates, so the optical density (OD) of each solution is directly proportional to the amount of purple formazan formed, which is itself a proxy of the number of viable cells.

Caco-2 and CCD-841 cells were seeded in 96-well plates (surface area per well = 0.29cm²) with 7.0×10^3 and 1.3×10^4 cells per well respectively, and grown until 70-80% confluent. Then cells were treated by adding to culture medium in 6 replicates, DMSO-diluted SFN (0.4 – 200.0µM) or DMSO (control); final DMSO concentrations were 0.05% (v/v). At 24 or 48 h following the addition of SFN, culture medium was supplemented with MTT (100µL of MTT per mL of culture medium) then cells were incubated for a further 1 h for MTT metabolism. Subsequently, medium was removed and purple formazan precipitates were dissolved in DMSO (100µL/well). OD₅₇₀ and OD₆₇₀ of the solutions in each well were measured using the FLUORstar Omega microplate reader (BMG Labtech).

$$\% \text{ Cell viability} = \frac{OD_{570}(\text{sample}) - OD_{670}(\text{sample})}{OD_{570}(\text{control}) - OD_{670}(\text{control})} \times 100$$

Data were analysed with the CalcuSyn software (Biosoft, Cambridge, UK) to determine IC₅₀ values.

2.11 Wound Healing Assays to Assess Cell Migration

Cells were seeded in 12-well plates (surface area/well = 3.8cm²) at densities such that cells would be 90-100% confluent at 24 h after seeding. At confluence, a vertical scratch wound was created down the centre of each monolayer using a 1 mL pipette tip. Cells were then washed in culture medium, and then treated by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control); final DMSO concentrations were 0.05% (v/v).

Following the addition of SFN, cells continued to be cultured under normal conditions and observed at 24 h intervals, until differences in wound closure between control and treated monolayers were noticeable by eye. Then, medium was removed and cells were washed in PBS before being fixed by incubation under 4% (v/v) PFA for 10 min. Fixed cells were washed in PBS and then stained by incubating under 0.1% (v/v) Crystal Violet solution (in water) for 30 min. Stained cells were washed in water and then photographed immediately or put into storage under PBS at 4°C.

Five photographs were taken along the vertical axis of each scratch wound, through an inverted microscope. Average gap widths between leading cells of scratch wounds in every image were determined using the ImageJ software, by averaging the measured gap widths at 8 fixed loci down the vertical axes of the scratch wounds. Averages across the five photographs of each scratch wound were then taken.

2.12 Statistics

Where $n \geq 3$, statistical comparisons were made using unpaired two-tailed Student's T-tests, which were carried out using the Microsoft Excel spreadsheet software. Results are expressed as means \pm S.E.M.

2.13 Rationale for Choice of Methods

2.13.1 RNA Extraction Methods

There are several types of method for the extraction of RNA from cultured cells, including organic extraction, filter-based extraction, magnetic particle-based extraction, and direct lysis.

2.13.1.1 Organic Extraction

Organic extraction involves lysing cells in a phenol-containing solution, then phase-separating the RNA from DNA and proteins by mixing in chloroform and centrifuging. RNA is isolated from aqueous phases, typically by alcohol precipitation. Such methods reliably maintain RNA integrity, and efficiently recover RNA species of all sizes. However, they tend to be more time-consuming than filter-based extraction methods, and are more hazardous due to the requirement of phenol and chloroform.

2.13.1.2 Filter-Based Extraction Methods

Filter-based extraction methods involve firstly lysing cells in a buffer supplemented with a strong inactivator of RNAses, such as 2-mercaptoethanol, and secondly, isolating RNA from the lysates by capture on glass fibre or silicon membranes in columns. Subsequent steps include washing the bound RNA with alcohol-based solutions to remove impurities, then eluting the RNA in water or an aqueous buffer. Extraction methods of this type tend to be much quicker than organic extraction methods, and avoid the chemical hazards of phenol and chloroform. However, most filter-based methods cannot efficiently recover RNA species of all sizes at once, but kits are generally optimised or optimisable for the efficient extraction of either small or large RNA species.

2.13.1.3 Direct Lysis

A simple approach to extracting RNA is to simply lyse cells in a solution that is able to effectively stabilise nucleic acids and that is compatible with the intended downstream applications. While this is fast and inexpensive, the presence in the lysates of other cellular macromolecules such as DNA and proteins makes this approach unsuitable for the majority of downstream

RNA-based assays, as well as for RNA concentration determination by spectrophotometry.

2.13.1.4 Choices of Methods

All intended downstream applications required RNA to be separated from DNA and proteins, so direct lysis was disregarded as an option.

The total RNA used for miRNA library construction was extracted by the organic guanidine thiocyanate-phenol-chloroform method, because it was necessary to effectively recover small RNA species such as miRNAs, whilst simultaneously recovering larger species in order to be able to evaluate overall RNA integrity via agarose gel electrophoresis and visualisation of 28S/18S bands. The total RNA used for RT-qPCR mRNA transcript assays was extracted using the filter column-based GenElute Total Mammalian RNA Kit (Sigma), which was considerably quicker than organic extraction, and it was not necessary to capture smaller RNA species in this context. For Northern Blotting, either total RNA was extracted with by organic guanidine thiocyanate-phenol-chloroform extraction, or small RNA fractions were extracted using the filter column-based miRPremier MicroRNA Isolation Kit (Sigma). The latter method is not able to efficiently recover large and small RNA species in the same fractions, but this was considered unnecessary for these assays, and small RNA extraction with the filter column-based kit was much quicker than organic total RNA extraction.

2.13.2 Protein Concentration Determination

Protein concentrations were determined using the Brilliant Blue G dye-binding assay of Bradford – a colorimetric assay based on the conversion of the Coomassie Brilliant Blue G-250 red dye to a blue compound upon reaction with protein. The dye donates electrons to ionisable groups on amino acid residues, which alters protein conformation in a manner that exposes hydrophobic pockets. The exposed hydrophobic pockets interact with the non-polar regions of the dye molecules, enabling the positively charged amine groups of amino acid residues to bind negatively charged regions of the dye molecules, which stabilises the dye in the blue form. While this assay is slightly less precise than some other methods of protein

concentration determination, such as the Lowry assay, it is significantly quicker and cheaper than such, and sufficiently precise for the intended application.

2.13.3 Protein Expression Assay Methods

There are several methods by which the relative abundance of specific proteins across samples can be compared, including Western Blotting, and the enzyme-linked immunosorbent assay (ELISA). In Western Blotting, proteins are denatured and size-fractionated by SDS-PAGE, transferred to a membrane, then incubated with antibodies specific to the proteins of interest, the binding of which is measured by various methods including those of a fluorescent or chemiluminescent nature. While it is considered to be slightly less precise than ELISA, it has several advantages over ELISA such as allowing the visualisation of protein degradation, and the assay of several proteins at once. Thus, Western Blotting was chosen, particularly since the aim was to detect significant, rather than subtle, differences.

2.13.4 Measurement of Cell Viability

There are several methods by which the effects of compounds on cell viability can be assayed, one of which is the MTT assay. There are several caveats to the use of the MTT assay to assess cell viability, such as the fact that the rate of MTT reduction by viable cells is not strictly consistent and that MTT has the potential to be reduced by superoxide (Wang, Yu et al. 2011), levels of which might be artificially altered by cell treatments. However, it is relatively fast and inexpensive, and should be sufficiently precise for the intended application.

Chapter 3: Establishing and Validating the Experimental System

3.1 General Introduction

SFN is known to be indiscriminately cytostatic and/or cytotoxic at very high doses, probably due to acute GSH depletion and/or DNA damage induction. Thus, the sub-cytotoxic range of concentrations suitable for experimental Caco-2 and CCD-841 treatment had to be determined. Therefore, both cell lines were treated with SFN across a range of concentrations, and effects on cell viability were assayed. Subsequently, the effects of SFN on two *in vitro* oncogenic 'endpoints': cell migration and clonogenicity, were assayed.

As discussed in Chapter 1, it has long been established that sub-cytotoxic doses of SFN are able to activate Nrf2 and thus upregulate the expression of various oxidant-reducing Nrf2-driven genes, across a range of cell lines and contexts. However, of relatively novel interest are the potential roles of miRNAs in the anti-cancer effects of SFN at sub-cytotoxic doses. It was necessary to first validate the suitability of the experimental set-up conditions, such as the cell density at which SFN was administered, the mode and vehicle of SFN delivery, and the time points following SFN treatment at which effects were assayed. Such validation was performed by assaying the effects of SFN treatment on both nuclear Nrf2 accumulation in Caco-2, and the expression of Nrf2-driven genes at the mRNA and/or protein levels in both cell lines. If the expected inductions were observed, similar conditions were to be maintained for later miRNA profiling experiments.

It is known that many miRNAs can directly and/or indirectly modulate Nrf2 activity. For example, several miRNAs of the let-7 family have been demonstrated to inhibit Bach1, which itself is able to repress Nrf2-mediated transcriptional activation by competing with Nrf2 for sMAF binding in the nucleus (Hou, Tian et al. 2012). Experimental overexpression of the let-7 miRNAs has been demonstrated to upregulate HO-1 and thereby reduce oxidative stress in human hepatocytes (Hou, Tian et al. 2012).

3.2 Effects of Sulforaphane and Hydrogen Peroxide on Cell Viability in Caco-2 and CCD-841

3.2.1 Sulforaphane

The MTT assay was used to measure relative cell viability at 24 or 48 h following the treatment of cells with SFN at various concentrations.

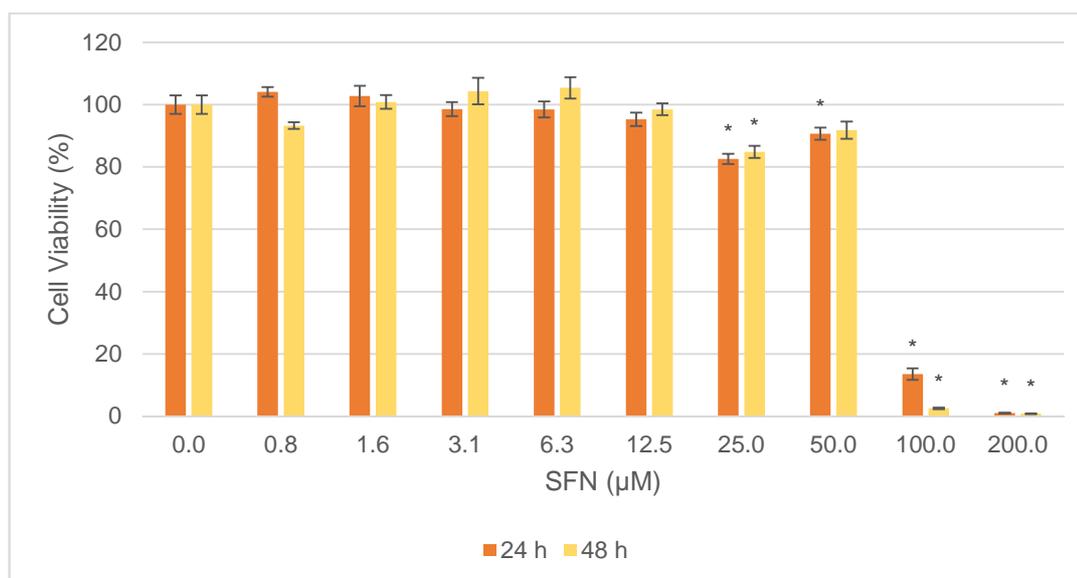


Figure 14: The effects of SFN on Caco-2 cell viability at 24 h and 48 h. Caco-2 cells were treated for 24 or 48 h, by adding to culture medium in sextuplicate, DMSO-diluted SFN at various concentrations, or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Cell viability was measured by the MTT assay and evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$).

As illustrated in Figure 14, SFN significantly reduced Caco-2 cell viability at concentrations $\geq 25\mu\text{M}$, at both 24 and 48 h, at which the half maximal inhibitory concentration (IC_{50}) values were very similar: 40 and 42 μM respectively. SFN cytotoxicity is not observable at concentrations $\leq 12.5\mu\text{M}$, implying that 10 μM should be an appropriate concentration for experimental use.

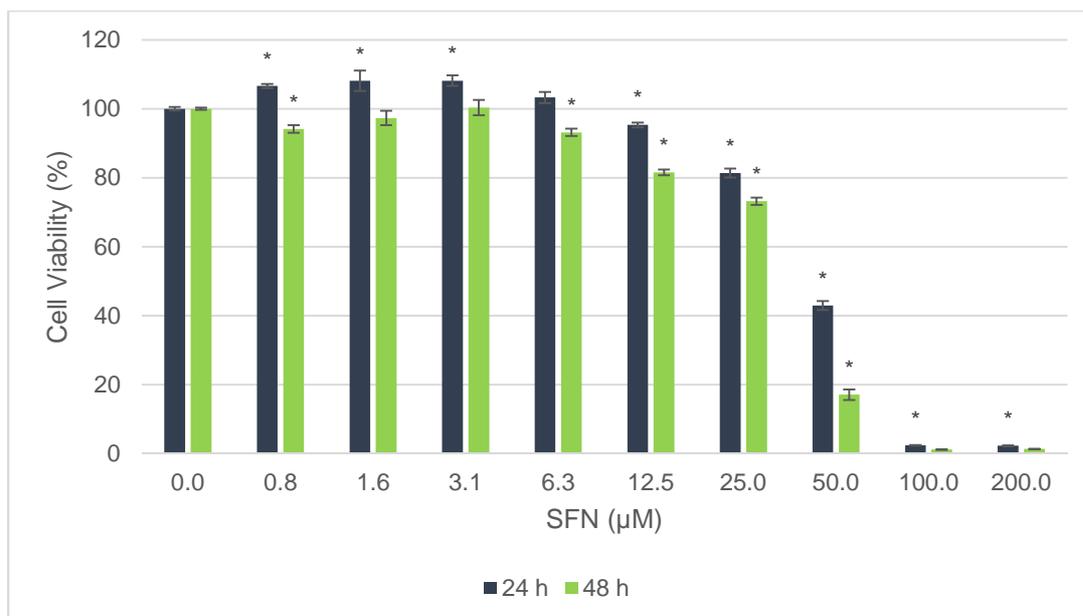


Figure 15: The effect of SFN on CCD-841 cell viability at 24 and 48 h. CCD-841 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, DMSO-diluted SFN at various concentrations, or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Cell viability was measured by the MTT assay and evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$).

As illustrated in Figure 15, CCD-841 cell viability at 24 h appeared to be increased by SFN at concentrations $\leq 3.13 \mu\text{M}$, but reduced at those $\geq 12.5 \mu\text{M}$. The concentrations of SFN that apparently increased viability at 24 h had either a negligible or inhibitory effect at 48 h, while those inhibitory at 24 h were even more so at 48 h. This could reflect a temporary promotion of proliferation at 24 h, which Nrf2 itself has been shown to induce in several cell lines (Homma, Ishii et al. 2009), (Gan, Sun et al. 2010), including those of a colonic nature (Arlt, Bauer et al. 2009). Considering that SFN probably induces pro- and anti-proliferative pathways simultaneously, it could be imagined that there exists a point in time between 24 and 48 h at which SFN's net effects on proliferation switch from positive to negative. The acute oxidant effects of SFN may activate proliferative ERK and MAPK signalling pathways. On the other hand, it is possible that the apparent increase in cell viability is just an artefact of the MTT assay. The MTT assay is based upon the reduction of tetrazolium salts, which can be reduced by superoxide

(Wang, Yu et al. 2011), and superoxide levels are probably modulated by SFN.

The IC_{50} [24 h] and IC_{50} [48 h] were 97 and 20 μ M respectively. Cytotoxicity is not evident at concentrations < 12.5 μ M, indicating that a concentration of 10 μ M should be experimentally appropriate.

3.2.2 Hydrogen Peroxide

The cytotoxicity of hydrogen peroxide (H_2O_2) (25 to 3200 μ M) – itself a source of ROS - towards Caco-2 and CCD-841 cells was assessed in a similar fashion to that of SFN.

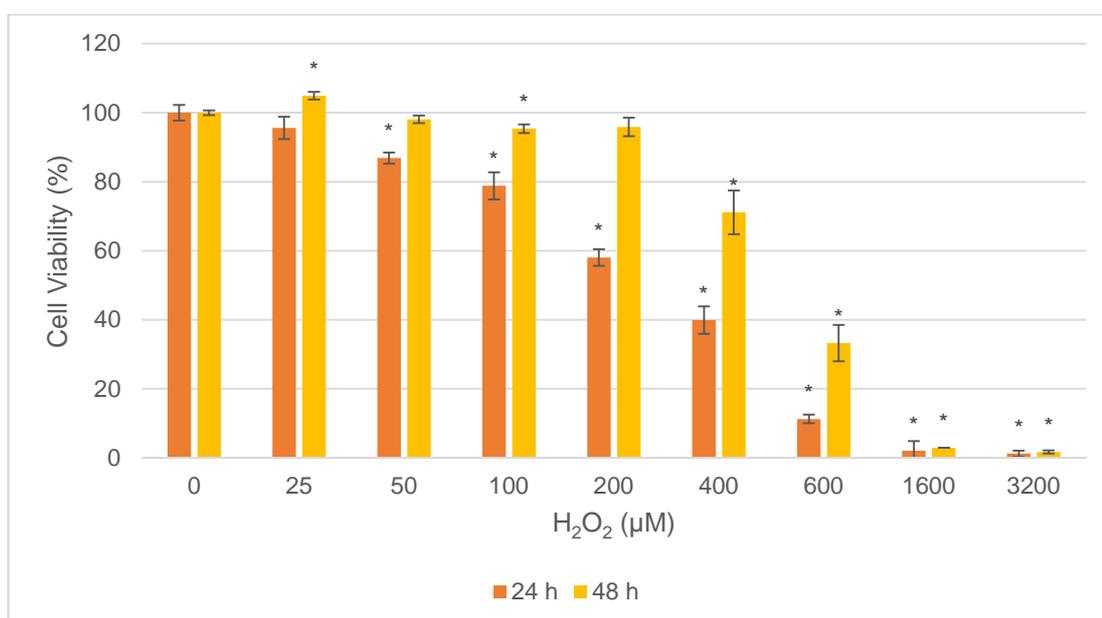


Figure 16: The effects of H_2O_2 on Caco-2 cell viability at 24 h and 48 h. Caco-2 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, water-dissolved H_2O_2 at various concentrations, or water (control). Cell viability was evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$).

As illustrated in Figure 16, H_2O_2 treatment inhibited Caco-2 cell viability at concentrations $\geq 50\mu$ M, albeit less so at 48 than at 24 h, indicating a partial recovery of cells between these two time points. IC_{50} [24 h] and IC_{50} [48 h] were 199 and 318 μ M respectively.

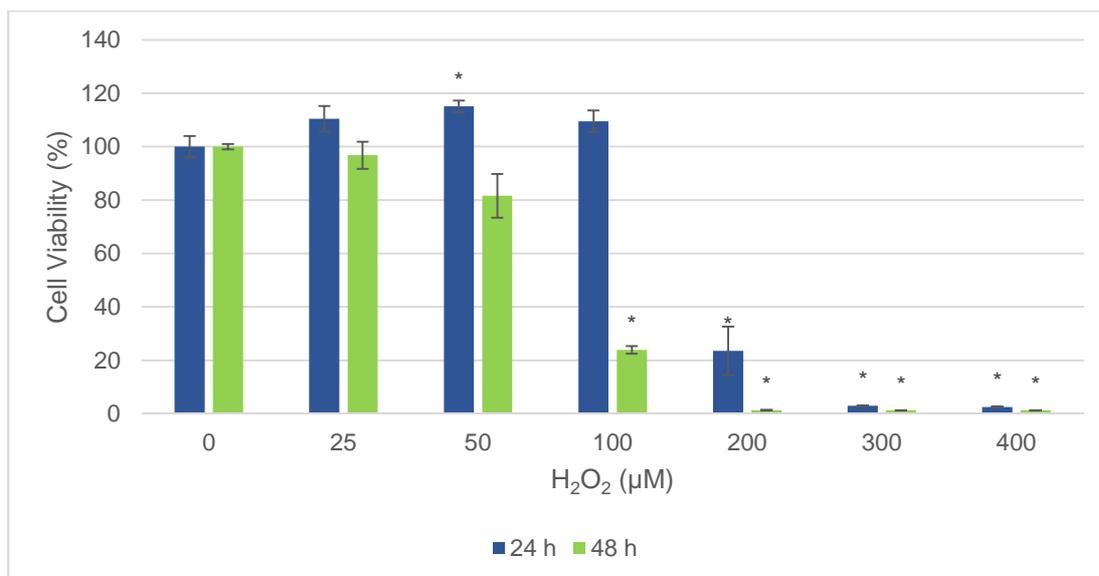


Figure 17: The effect of H₂O₂ on CCD-841 cell viability at 24 h and 48 h. CCD-841 cells were treated for 24 or 48 h by adding to culture medium in sextuplicate, water-dissolved H₂O₂ at various concentrations, or water (control). Cell viability was evaluated as (OD₆₇₀ – OD₅₇₀) at each concentration in proportion to the same of the control, and data are represented as the mean ± S.E.M. of sextuplicates (*P<0.05).

The data in Figure 17 indicate that H₂O₂ treatment significantly inhibited CCD-841 cell viability at concentrations ≥ 200μM at 24 h, and ≥100μM at 48 h. 50μM H₂O₂ apparently increased cell viability at 24 h but not at 48 h. IC₅₀[24 h] and IC₅₀[48 h] were 197 and 68μM respectively.

Unlike Caco-2 cells, CCD-841 cells do not appear to recover between 24 and 48 h, implying greater sensitivity to H₂O₂-mediated cytotoxicity and/or the existence of delayed cytotoxic pathway induction. A possible explanation could be higher basal activity of p53 in CCD-841 vs. Caco-2, thus making the CCD-841 cells more sensitive to H₂O₂-induced cell cycle arrest and/or apoptosis. The gene coding for p53, *TP53*, is mutated in several CRC cell lines (Rodrigues, Rowan et al. 1990); its inactivation is considered to be the final critical step to metastasis in the Vogelstein model of pathogenesis (Fearon and Vogelstein 1990).

It is also possible that H₂O₂ upregulates proliferative and anti-apoptotic pathways in Caco-2 via the activation of PKC proteins, upon binding to their regulatory domains (Gopalakrishna and Gundimeda 2002).

3.2.3 Sulforaphane vs. Hydrogen Peroxide

The cell viability assays indicate that CCD-841 cells were much more sensitive than Caco-2 to H₂O₂; the IC₅₀ of H₂O₂ at 48 h was 68µM in CCD-841 cells vs. 318µM in Caco-2. Caco-2 cells appeared to 'recover' from H₂O₂ significantly between 24 and 48 h, at which IC₅₀ values were 199 and 318µM respectively. However, they did not similarly 'recover' from SFN treatment between 24 and 48 h, at which IC₅₀ values were very similar at 40 and 42µM respectively. This may imply roles for oxidative stress-independent mechanisms in the suppression of Caco-2 cell proliferation by SFN.

3.2.4 Methylselenocysteine (SeMSC)

Methylselenocysteine (SeMSC) is a supplemental source of selenium, which is vital for the synthesis and activity of important cytoprotective reducing (antioxidant) enzymes such as TrxR-1. Selenoproteins, such as TrxR-1, incorporate a selenocysteine amino acid residue that is coded for by an alternative reading of the UGA 'STOP' codon. This alternative reading is initiated by the presence of a stem-loop-forming motif within the 3'-UTR, called the selenocysteine insertion sequence, and requires the presence of specific binding proteins in addition to a specific tRNA for selenocysteine (Mathers and Hesketh 2007).

The effects of SeMSC on cell viability were assayed over a range of concentrations using the MTT assay, the data from which are shown in Figure 18.

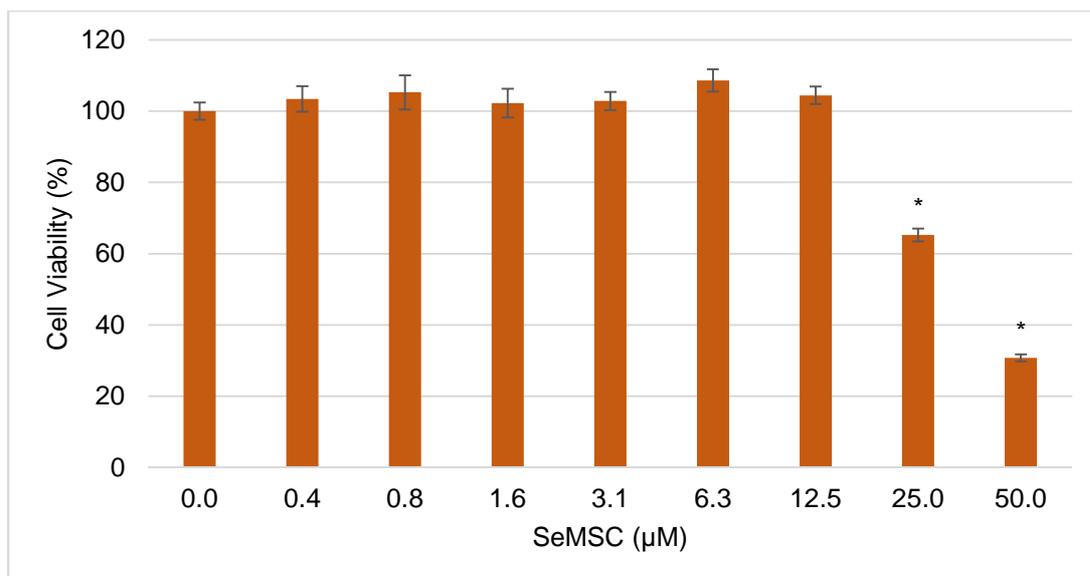


Figure 18: The effects of SeMSC on Caco-2 cell viability at 24 h. Caco-2 cells were treated for 24 h by adding to culture medium in sextuplicate, water-dissolved SeMSC at various concentrations, or water (control). Cell viability was evaluated as $(OD_{670} - OD_{570})$ at each concentration in proportion to the same of the control, and data are represented as the mean \pm S.E.M. of sextuplicates (* $P < 0.05$).

3.3 Effects of Sulforaphane on Nrf2 and Nrf2-Controlled Phase II Enzyme Expression in Caco-2 and CCD-841 Cells

3.3.1 Thioredoxin Reductase 1

TrxR-1 is a phase II reductant (antioxidant) enzyme whose gene promoter contains the *ARE*, thus making it liable to Nrf2-mediated transcriptional activation.

3.3.1.1 Messenger RNA Level

The effects of 24 h SFN treatment on TrxR-1 at the mRNA level, in both Caco-2 and CCD-841 cells, were assayed by TaqMan RT-qPCR. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (at concentrations of 2.5, 5, 10 or 20 µM), or DMSO (control); final DMSO concentrations were 0.1% (v/v).

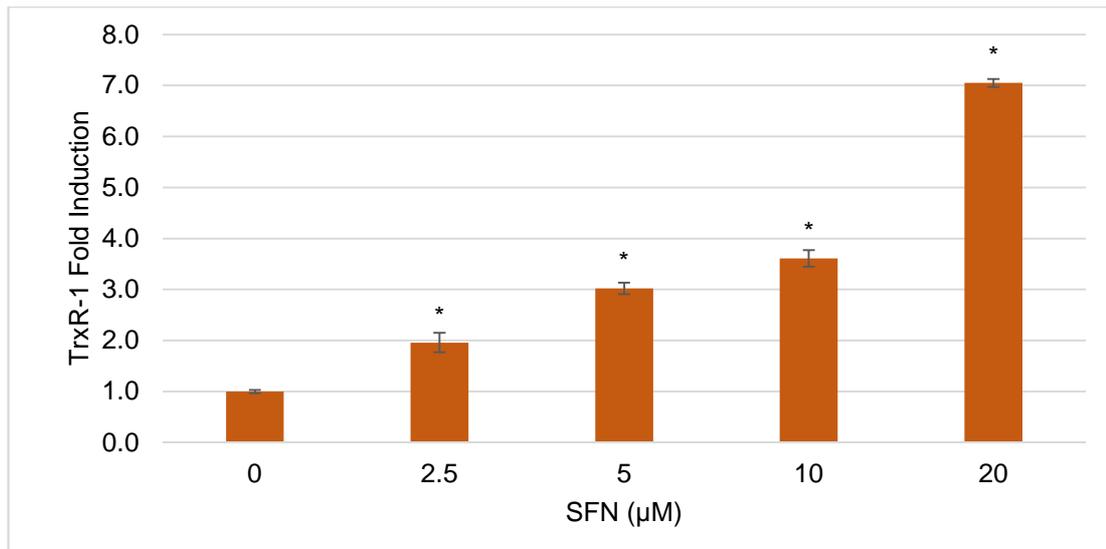


Figure 19: The effects of 24 h SFN treatment on the TrxR-1 mRNA transcript level in Caco-2 cells. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (2.5, 5, 10 or 20μM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Relative TrxR-1 mRNA levels were assayed via TaqMan RT-qPCR as described in Methods, using 18S as an internal control. Relative TrxR-1 expression is evaluated by the Pfaffl method from C_T values, taking into account efficiencies derived from standard curve data. Data are represented as mean \pm S.E.M. of triplicates (* $P < 0.05$).

As shown by the RT-qPCR data illustrated in Figure 19, TrxR-1 was clearly and dose-dependently upregulated in Caco-2 at the mRNA level by 24 h SFN (2.5 to 20μM) treatment.

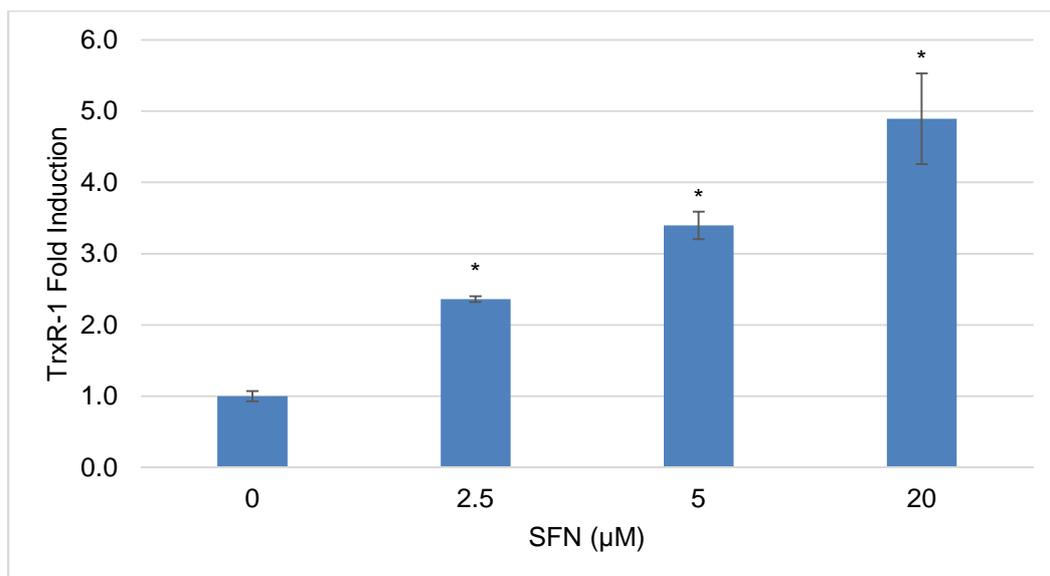


Figure 20: The effects of 24 h SFN treatment on the TrxR-1 mRNA transcript level in CCD-841 cells. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (2.5, 5, 10 or 20μM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Relative TrxR-1 mRNA levels were assayed via TaqMan RT-qPCR as described in Methods, using 18S as an internal control. Relative TrxR-1 expression is evaluated by the Pfaffl method from C_T values, taking into account efficiencies derived from standard curve data. Data are represented as mean ± S.E.M. of triplicates (* $P < 0.05$).

3.3.1.2 Protein Level With/Without Supplemental Selenium

As is clear from the RT-qPCR data illustrated in Figure 19 and Figure 20, TrxR-1 was dose-dependently upregulated at the mRNA level at 24 h in both Caco-2 and CCD-841 cells, so effects of SFN at the protein level were then tested.

As previously mentioned, TrxR-1 is a selenoprotein i.e. it contains selenocysteine residues that are coded for by the UGA codon, which is ordinarily a STOP codon, but can alternatively code for selenocysteine in the presence of the selenocysteine insertion sequence in the 3'-UTR and Sec-tRNA^{Sec}. Initially, the selenocysteine tRNAs are actually charged with serine to form Ser-tRNA^{Sec}, whose serine residue can be converted to selenocysteine by selenocysteine synthase to form the required Sec-tRNA^{Sec}. This enzyme transfers the selenium from selenium phosphate, which itself is formed from selenide and ATP by selenophosphate synthetase 1 (Low, Harney et al. 1995). Cells store selenium in the form of selenide.

Under selenium deficiency, the absence of selenide means that selenophosphate cannot be formed, and thus neither can Ser-tRNA^{Sec} be converted to Sec-tRNA^{Sec}. Ser-tRNA^{Sec} itself cannot be incorporated into the peptide chain because it is not recognised by the translation elongation factor eEF1A, so translation stops in the absence of Sec-tRNA^{Sec} (Xu, Carlson et al. 2007). For this reason, the full translation of TrxR-1 can be inhibited by selenium deficiency.

Western Blotting was used to assay the effects of SFN on TrxR-1 at the protein level in both cells, either by treatment with SFN at different concentrations (2.5 to 20 μ M) for 48 h, or by treatment for different lengths of time (6 to 24 h) with a fixed concentration of 10 μ M. SFN treatment was performed in both the presence and absence of supplemental SeMSC.

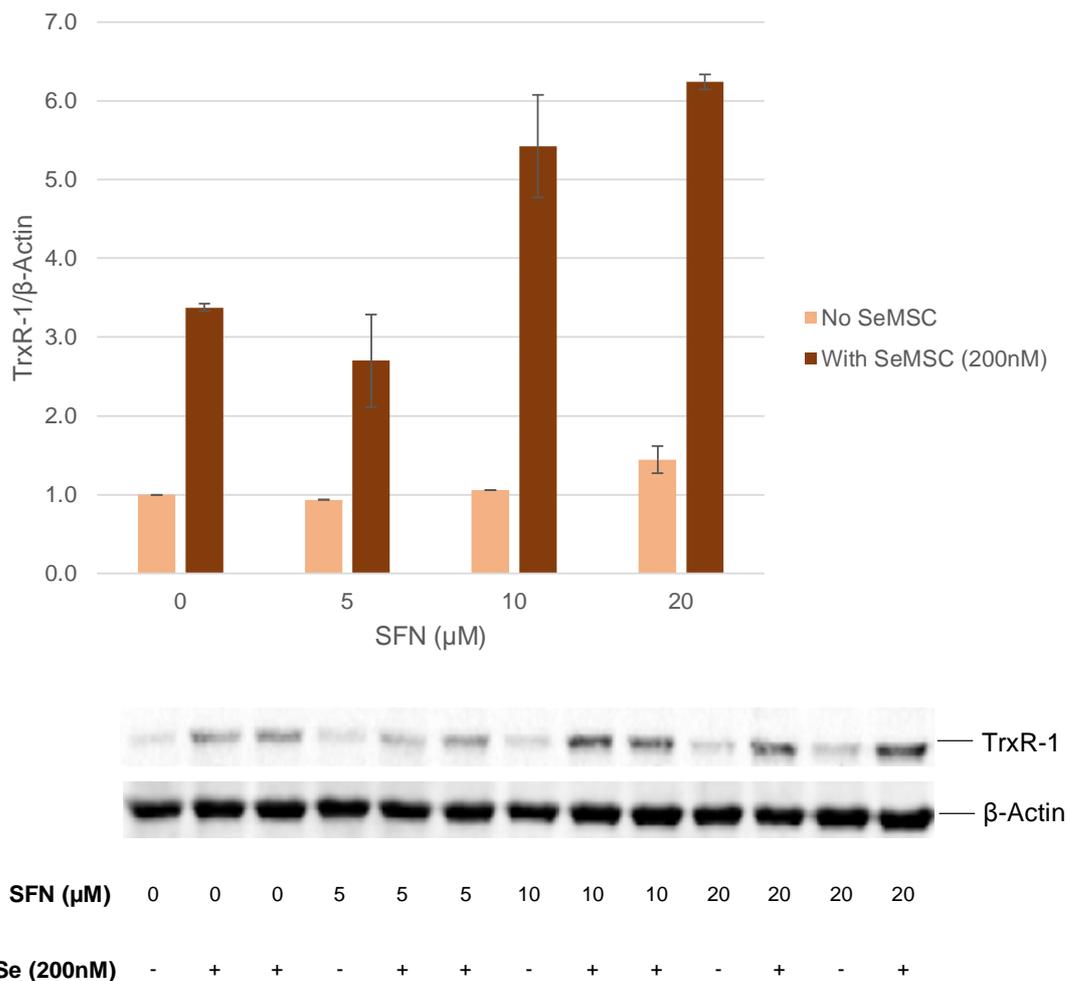


Figure 21: The effect of 48 h SFN treatment on TrxR-1 protein levels in Caco-2 cells. Cells were treated for 48 h by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20μM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2.

From the data illustrated in Figure 19, it is clear that TrxR-1 is dose-dependently upregulated at the mRNA level by SFN treatment, without the need for selenium supplementation. However, those illustrated in Figure 21 and Figure 22 demonstrate that corresponding upregulation at the protein level is largely dependent upon SeMSC supplementation.

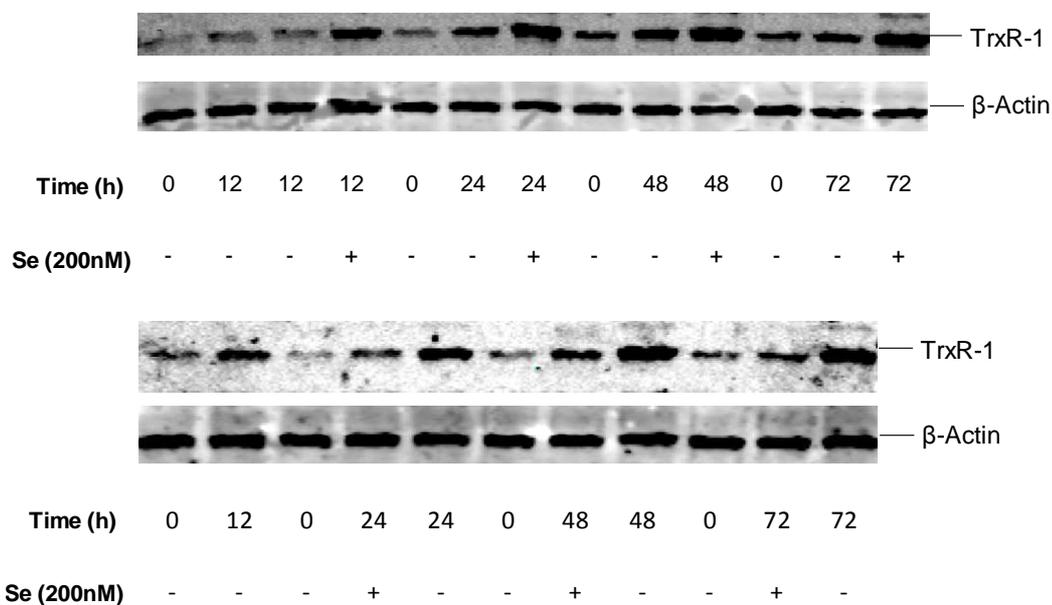
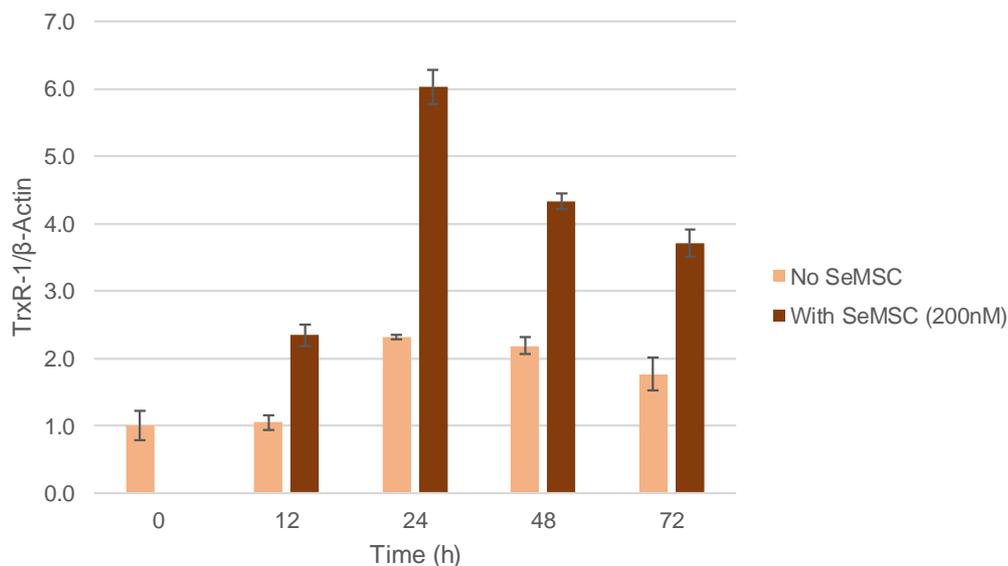


Figure 22: The effect of SFN treatment on TrxR-1 protein levels in Caco-2 cells, with or without selenium supplementation. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM). Relative protein abundance was assayed via Western Blotting as described in Methods, using β -actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2.

The Western Blot data illustrated in Figure 22 demonstrate that following the treatment of Caco-2 cells with SFN (10 μ M) and SeMSC (200nM), TrxR-1

protein levels peaked at 24 h, before declining at 48 h, albeit to levels still significantly above baseline, which were apparently maintained up to 72 h.

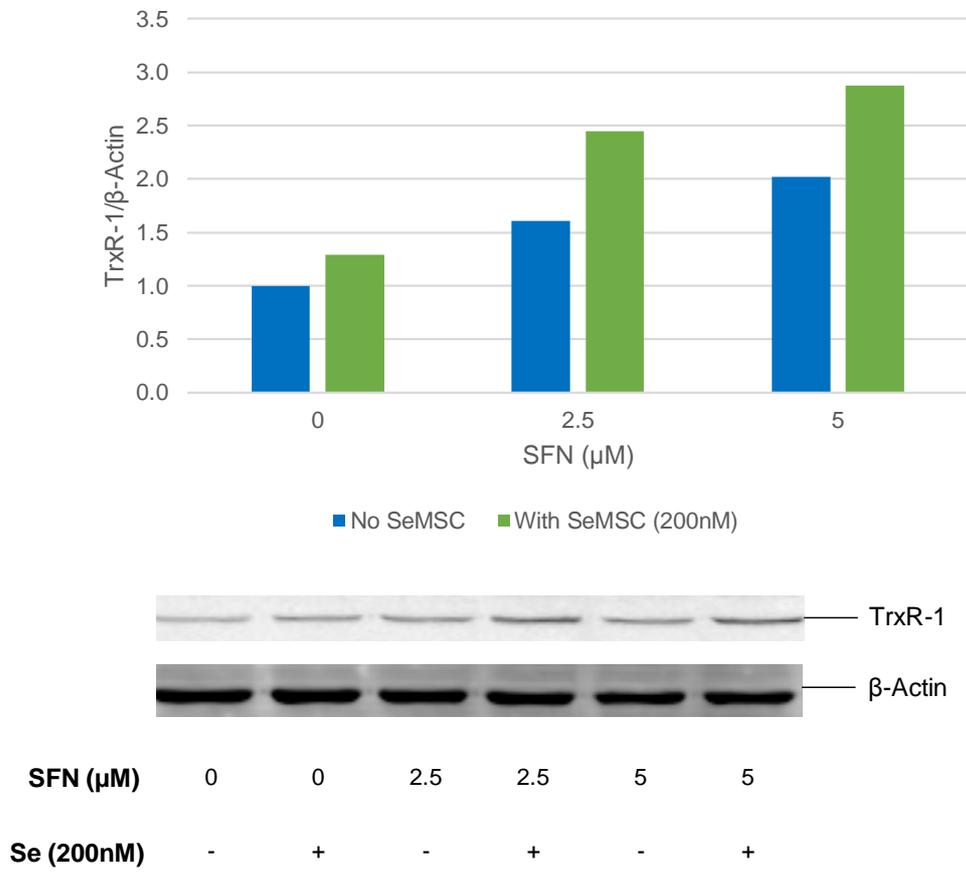
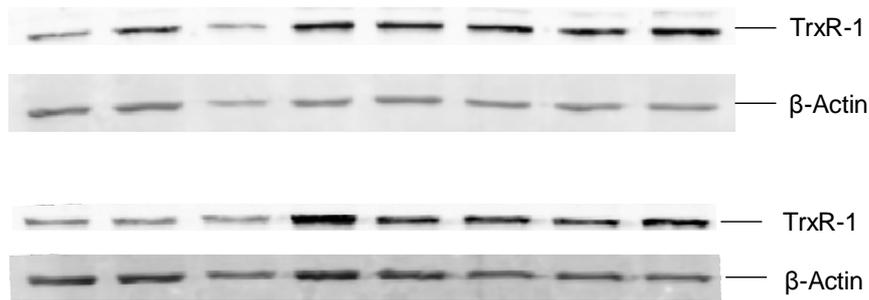
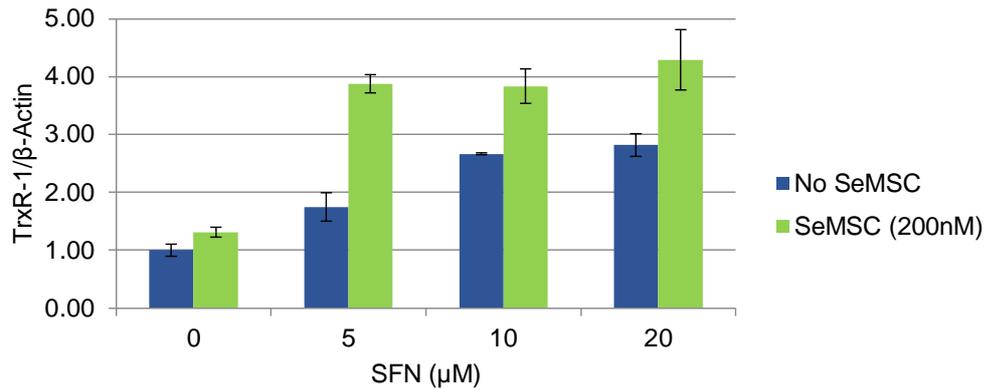


Figure 23: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated for 24 h by adding to culture medium, DMSO-diluted SFN (2.5 or 5μM) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are representative of a single experiment.



SFN (μM)	0	0	5	5	10	10	20	20
Se (200nM)	-	+	-	+	-	+	-	+

Figure 24: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated for 48 h by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20μM) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM), to culture medium. Relative protein levels were assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates ± (upper value – lower value)/2.

As illustrated by the Western Blot data in Figure 23, 24 h SFN and SeMSC treatment synergistically upregulated the TrxR-1 protein in CCD-841, as did 48 h treatment of the same as illustrated by the data in Figure 24.

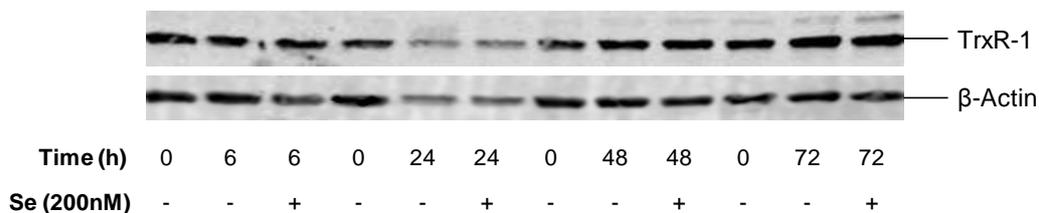
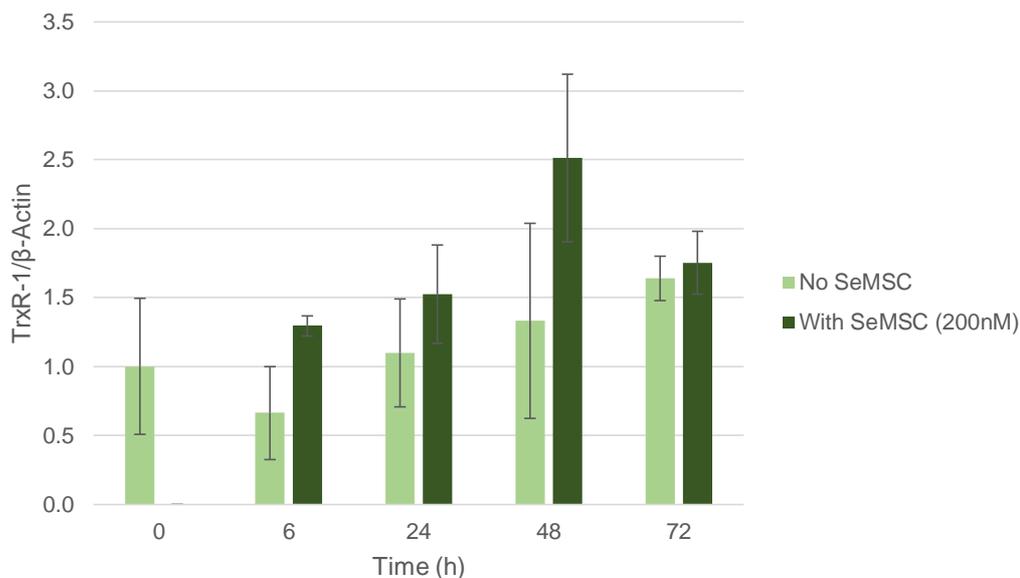


Figure 25: The effect of SFN treatment on TrxR-1 protein levels in CCD-841 cells, with or without selenium supplementation. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) or DMSO (control) (final DMSO concentrations were 0.1% (v/v)). Selenium was supplemented by adding SeMSC (200nM) to culture medium. Relative protein abundance was assayed via Western Blotting as described in Methods, using β -actin as an internal control. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2.

In the presence of supplemental SeMSC, SFN (10 μ M)-mediated induction of TrxR-1 at the protein level peaked at 48 h, as illustrated by the data in Figure 25. At 72 h, TrxR-1 protein levels appear to be modestly elevated above baseline in response to SFN treatment, irrespective of selenium supplementation.

Caco-2 vs. CCD-841 Cells

In the presence of supplemental selenium, the induction of TrxR-1 in Caco-2 is apparently dose-dependent at concentrations $\geq 10\mu$ M. In CCD-841,

induction at 24 h is slightly greater following treatment with a concentration of 5 μ M vs. that following such with a concentration of 2.5 μ M, but induction at 48 h does not significantly increase as the concentration is raised above 5 μ M. Perhaps factors other than selenium availability limit TrxR-1 translation in this context.

3.3.2 UDP Glucuronosyltransferase 1 Family, Polypeptide A Cluster

UGT1A is a phase II reducing (antioxidant) enzyme, whose gene promoter contains the *ARE*, the binding of Nrf2 to which activates the gene's transcription. The effects of 24 h SFN (5 to 20 μ M) treatment on UGT1A expression in Caco-2 at the protein level were assayed by Western Blotting.

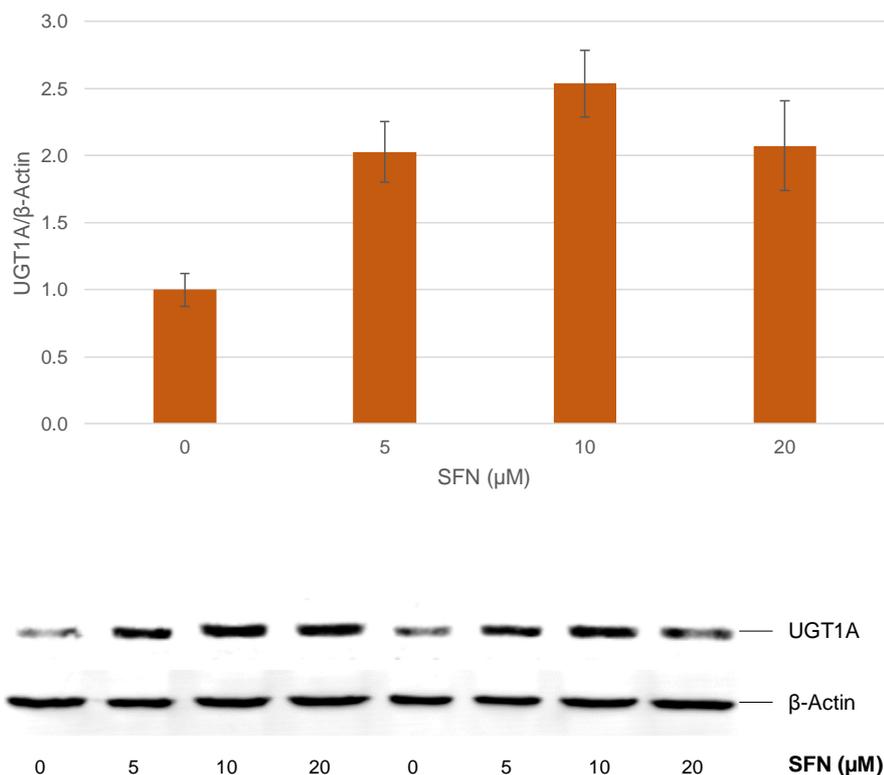


Figure 26: The effect of 24 h SFN treatment on UGT1A protein levels in Caco-2 cells. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Relative protein abundance was assayed via Western Blotting as described in Methods, using β -actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each

sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2.

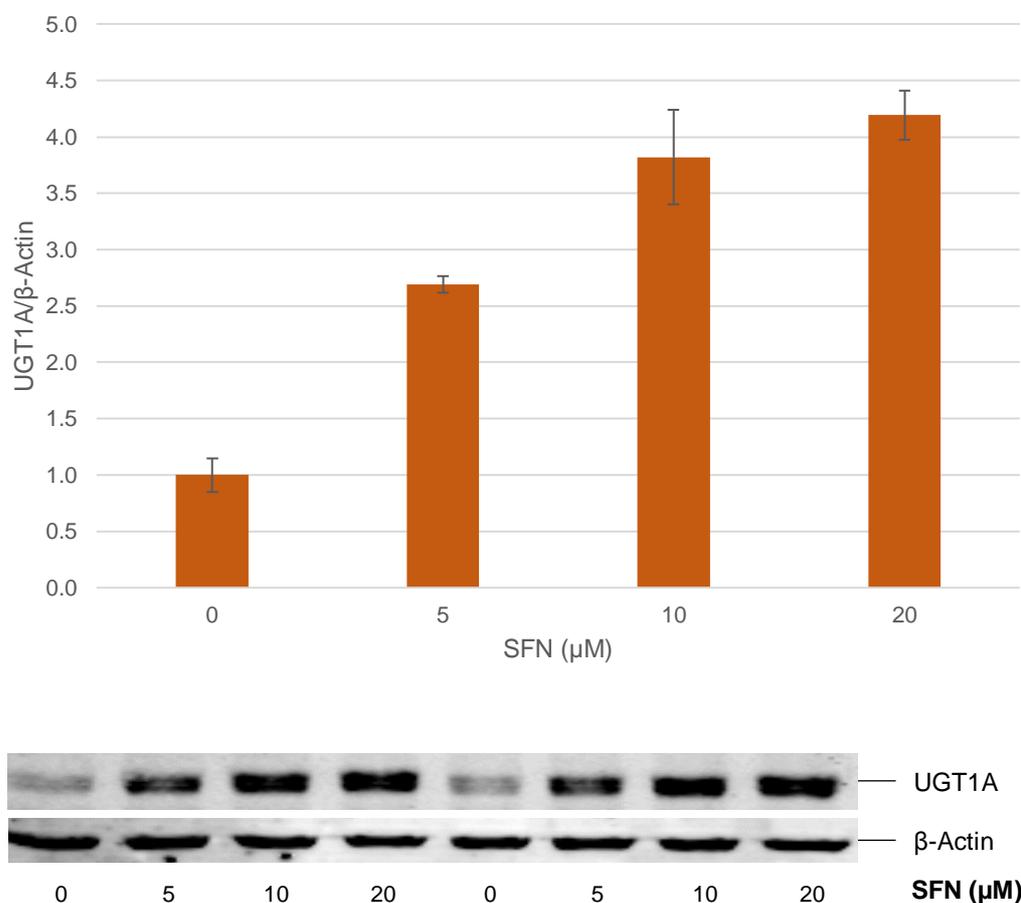


Figure 27: The effect of 72 h SFN treatment on UGT1A protein levels in Caco-2 cells. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20 μM) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Relative protein abundance was assayed via Western Blotting as described in Methods, using β-actin as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(TrxR-1 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2.

The Western Blot data represented in Figure 26 and Figure 27 clearly indicate that 24 and 72 h SFN (5-20 μM) treatment induced UGT1A in Caco-2 at the protein level. At 24 h, induction was dose-dependent at concentrations $\leq 10 \mu\text{M}$, but slightly lower at 20 than at 10 μM. At 72 h, however, induction of UGT1A was dose-dependent throughout the range of concentrations. Perhaps SFN simultaneously induces UGT1A-repressing and UGT1A-

stimulating pathways, with the latter gaining the upper hand in all cases tested except at 24 h following 20 μ M treatment.

3.3.3 Nrf2 Nuclear Accumulation

As previously discussed, Nrf2 induction involves the translocation of cytoplasmic Nrf2 to the nucleus, where it is able to exert activity. Western Blots were carried out with nuclear protein extracts from Caco-2 cells that had been treated with SFN across a 24 h time period, the data for which are shown in Figure 28.

It is apparent from the data illustrated in Figure 28 that nuclear Nrf2 levels are dramatically increased by SFN treatment from as early as 4 h, although not dose-dependently ($\geq 5\mu$ M). Nuclear Nrf2 levels appear to be highly variable, yet to remain significantly above baseline between 4 and 24 h post-treatment. The apparent dose dependence of TrxR-1 and UGT1A SFN-mediated induction may imply that factors other than Nrf2 translocation contribute to the upregulation of such Nrf2-controlled genes e.g. the inactivation of Bach1.

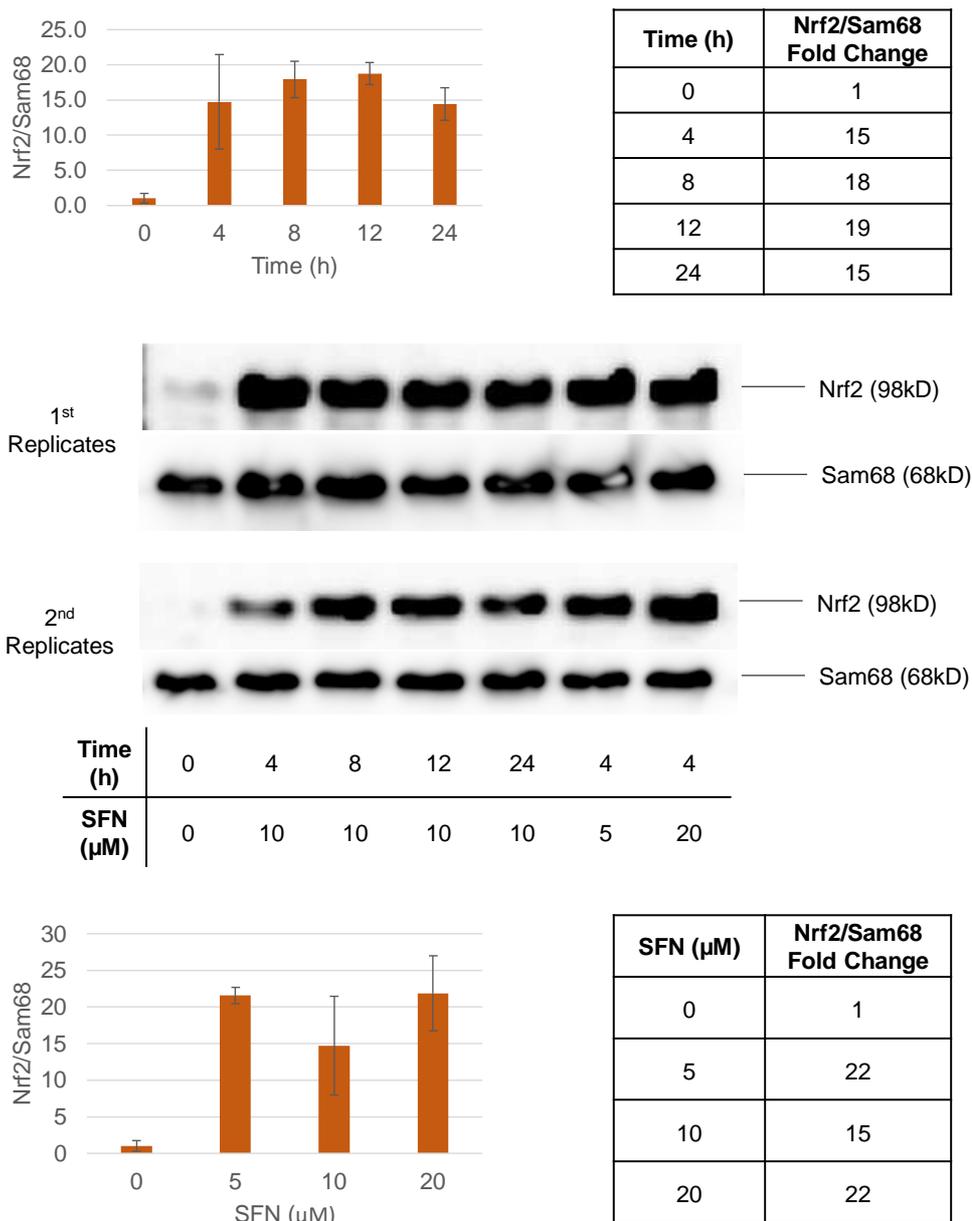


Figure 28: The effect of SFN treatment on nuclear accumulation of Nrf2 in Caco-2 cells. Cells were treated for 4, 8, 12 or 24 h with SFN (10 μ M) or for 4 h with SFN (5 or 20 μ M) by adding to culture medium in duplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control) (final DMSO concentrations were 0.05% (v/v)). Relative protein levels were assayed via Western Blotting as described in Methods, using Sam68 as an internal control for normalisation. Induction is evaluated and illustrated in the graph as [(Nrf2 band intensity)/(Internal control band intensity)] for each sample relative to the same for the control (see Methods). Data are represented as the mean of duplicates \pm (upper value – lower value)/2.

3.4 Effects of Sulforaphane on Oncogenic Characteristics

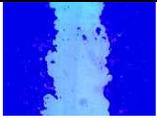
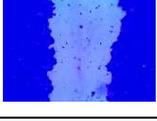
3.4.1 Effects on Cell Migration

The ability of cancer cells to migrate *in vitro* is considered to be a marker of oncogenicity, and thus agents that inhibit such migration are potentially tumour suppressive. The wound healing assay was used to evaluate the effects of SFN treatment on the migration of Caco-2 and CCD-841 cells across gaps created in confluent monolayers as described in Methods.

The data shown in Figure 29 provide no evidence that SFN has any effect on Caco-2 cell migration at 72 h following treatment.

By contrast, SFN significantly and dose-dependently inhibited the migration of CCD-841 cells over 48 h, as illustrated by the data presented in Figure 30.

It was surprising that the migration of the CCD-841 (non-cancerous) cells was inhibited while that of the Caco-2 (cancerous) cells was not. Perhaps Caco-2 cells have mutations or epimutations that confer resistance to the pathway(s) by which SFN (at 5-20 μ M) inhibits cell migration. However, the lack of effect on Caco-2 migration observed here does not necessarily prove that colonic adenocarcinoma cells are always resistant to SFN-induced migration repression. A possible confounding issue encountered with the Caco-2 cells was that it took longer than 24 h for all cells to attach following seeding. However, incubating the cells for longer than 24 h allowed strong intercellular adhesions to form, which resulted in the failure of attempts to create scratch wounds due to the entire monolayer peeling. Therefore, scratch wounds had to be created at 24 h. Inevitably, this selected for cells that attached more quickly, since those not attached by 24 h were discarded. Also, the highly adhesive nature of the Caco-2 cells may have meant that very little or no migration took place following wound creation, with or without SFN.

SFN (μM)	Replicate		
	1	2	3
0			
5			
10			
20			

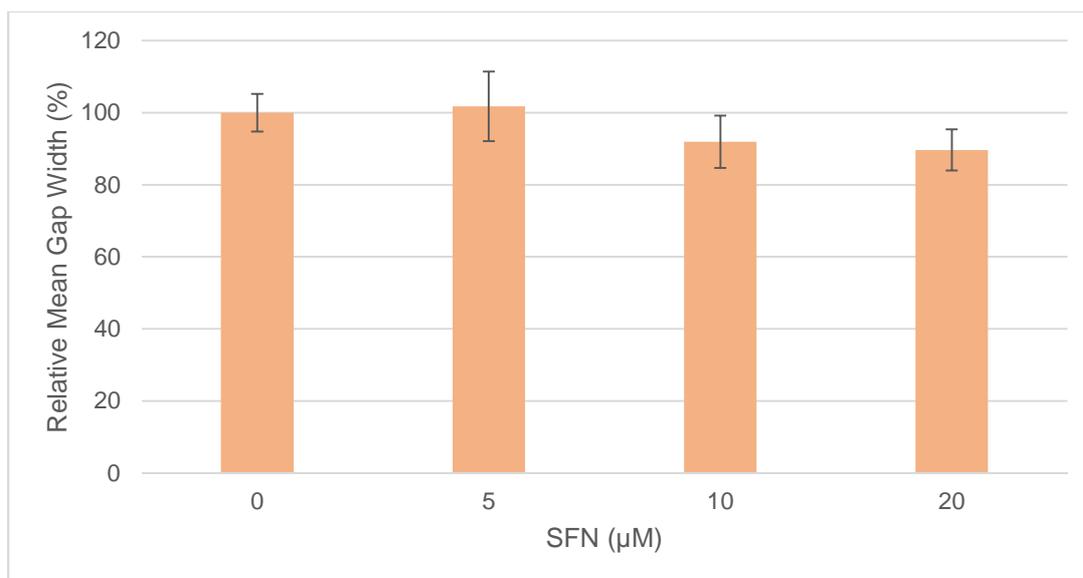
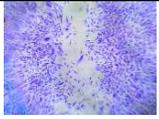
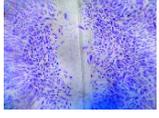
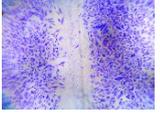
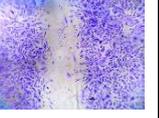
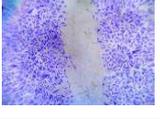
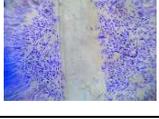
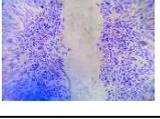


Figure 29: Caco-2 cells were seeded at densities sufficient that confluent monolayers were formed by 24 h. Scratch wounds were created in confluent monolayers using a pipette tip, and cells were then treated by adding to culture medium in triplicate SFN (5 to 20 μM) or DMSO; final DMSO concentrations were 0.05% (v/v). Cells were subsequently incubated for 72 h, then fixed, stained and photographed (with 5 photographs along the vertical axis of each scratch wound). The 4th image of every wound is shown. The average gap width between leading cells was calculated in each image, then data across all 5 images per well were averaged. Data from image analyses are represented in the graph as the means of three biological replicates \pm S.E.M.

The data for the assays in CCD-841 cells clearly indicate that SFN significantly inhibited wound healing in a dose-dependent manner - even at the lowest concentration (5 μ M) - indicating that SFN can inhibit cell migration independently of cytotoxicity. The implications of this for the effects of SFN on healthy colonic tissues are unclear, in terms of whether or not SFN could impair healing and/or renewal of the colonic epithelium. It might be postulated that the effects observed in this colonic cell line - CCD-841 – are likely to similarly occur in early-stage colonic tumour cells.

SFN (μM)	Replicate		
	1	2	3
0			
5			
10			
20			

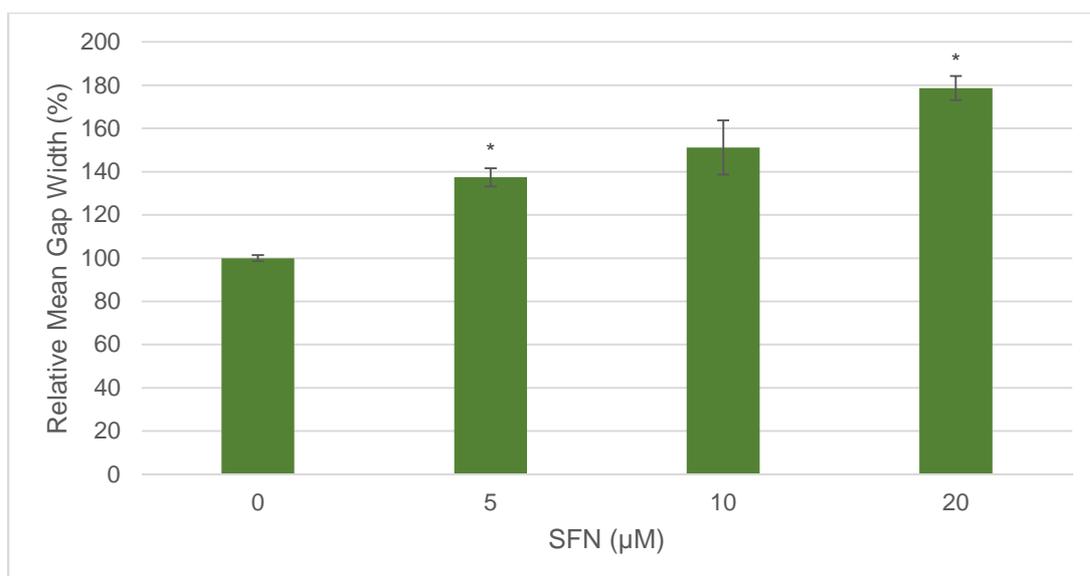


Figure 30: CCD-841 cells were seeded so as to form confluent monolayers at 24 h. Scratch wounds were then created in confluent monolayers with a pipette tip, and cells were then treated by adding to culture medium in triplicate SFN (5 to 20 μM) or DMSO; final DMSO concentrations were 0.05% (v/v). Cells were subsequently incubated for 48 h, then fixed, stained and photographed (with 5 photographs along the vertical axis of each scratch wound). The 4th image of every wound is shown. The average gap width between leading cells was calculated in each image, then data across all 5 images per well were averaged. Data from image analyses are represented in the graph as the means of three biological replicates \pm S.E.M. (* $P < 0.05$).

3.4.2 Effects on Clonogenicity

The impact of SFN (5-20 μ M) on the ability of Caco-2 cells to form colonies (≥ 50 cells) over several weeks after being seeded at very low density, i.e. cell clonogenicity, was tested. The potential of cells to form colonies following seeding at very low densities is regarded as a marker of oncogenicity *in vitro*, since it implies that the cells survive and undergo many cycles of division, as opposed to undergoing apoptosis and/or senescence, in spite of their isolation. The effects of SFN treatment on the clonogenicity of Caco-2 cells were assayed, as described in Methods.

The data shown in Figure 31 clearly show that the addition of SFN (5, 10 or 20 μ M) to culture medium after low-density seeding inhibited the 14-day colony-forming potential of Caco-2 cells in a dose-dependent fashion. However, it must be considered that since cell density was very low upon the addition of SFN, the cells were exposed more highly to SFN than if the same concentrations of SFN had been added to cells present at 70% confluence. Therefore, it cannot be ruled out that the inhibition of clonogenicity here observed was at least partly attributable to general cytotoxicity via oxidative and/or genotoxic stress. It is not known how rapidly SFN accumulated in cells following its addition to culture medium (Zhang 2000).

Subsequently, a similar colony forming assay was done but testing a different mode of SFN treatment i.e. the SFN pre-treatment of cells at 70-80% confluence prior to low-density seeding. Cells were cultured until 70-80% confluent in 24-well plates, and then treated by adding SFN/DMSO to culture medium for 24 h. Then the cells were detached using trypsin and re-seeded at low densities in 6-well plates. These cells were cultured as before, except that SFN was not present in the culture medium once the cells had been re-seeded at low density. Colonies were then stained and counted as before. The data from this second assay are summarised and illustrated in Figure 32.

SFN (μM)	Colony Count	% of Control Mean	Mean (%)
0	21	82	100
	21	82	
	25	98	
	26	102	
	28	110	
	32	125	
5	4	16	48
	7	27	
	8	31	
	12	47	
	21	82	
	22	86	
10	3	12	31
	5	20	
	6	24	
	9	35	
	11	43	
	13	51	
20	0	0	8
	0	0	
	2	8	
	3	12	
	3	12	
	5	20	

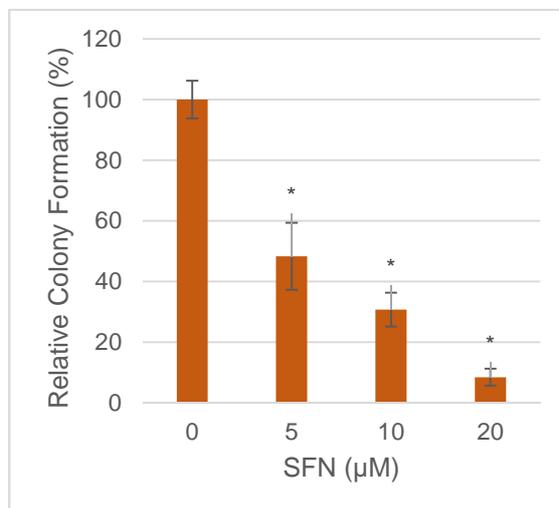


Figure 31: Caco-2 cells were seeded at low density (400 cells/well) in 6-well plates (surface area/well = 9.5cm²) and then incubated for 48 h prior to treatment. After 48 h, cells were treated by adding to culture medium in sextuplicate, DMSO-diluted SFN (5-20 μM) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Cells were then incubated for a further 14 days, after which cells were fixed and stained, and the number of colonies in each well was counted under an inverted microscope. Absolute colony counts are presented in the table and relative colony formation is illustrated in the graph, where data are represented as means of six biological replicates \pm S.E.M. (*P<0.05).

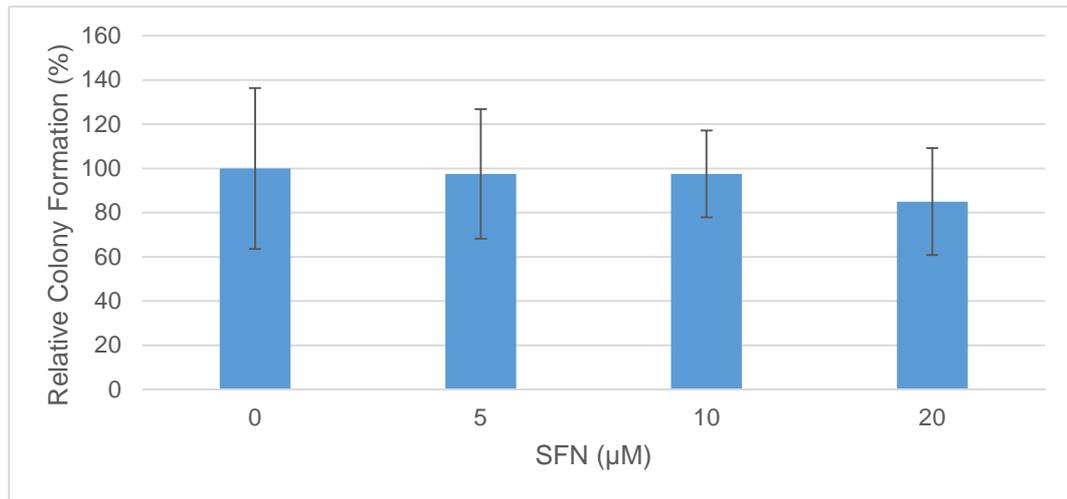


Figure 32: Caco-2 cells were cultured to 70-80% confluence in 24-well plates (surface area/well = 1.9cm²), then treated with DMSO-diluted SFN (5, 10 or 20μM) or DMSO; final DMSO concentrations were 0.05% (v/v). Subsequently, cells were detached using trypsin and then re-seeded at low density (400 cells/well) in 6-well plates (surface area/well = 9.5cm²). Cells were then incubated for a further 14 days after which cells were fixed and stained, and the number of colonies in each well was counted. Relative colony formation is illustrated in the graph, where data are represented as means of six biological replicates ± S.E.M.

It is apparent from the data illustrated in Figure 32 that the SFN pre-treatment of Caco-2 cells at 70-80% confluence prior to their detachment and re-seeding at low density did not affect clonogenicity. Perhaps single exposures to such sub-cytotoxic doses of SFN do not affect clonogenicity. However, this second assay did not monitor the effects of repeated sub-cytotoxic SFN exposure upon the clonogenicity of cells already in isolation, which may be more relevant from the perspective of tumour metastasis *in vivo*. It is also possible that effects were masked by cellular stress from the trypsin detachment and re-seeding process.

3.5 Conclusions

It is clear from the data presented in this Chapter that the tested experimental set-up is functional in terms of demonstrating the expected effects of SFN treatment on Nrf2 nuclear translocation and Nrf2-mediated gene expression in both Caco-2 and CCD-841 cells. Therefore, such set-up is appropriate for miRNA profiling experiments.

Chapter 4: Profiling MicroRNA Expression

4.1 General Introduction

The previous Chapter describes how the experimental system was validated by demonstrating the ability of sub-cytotoxic doses of SFN to induce the nuclear translocation of Nrf2, and upregulate the expression of Nrf2-controlled genes such as TrxR-1 in Caco-2 and CCD-841 cells, at the transcriptional and/or translational levels. Thus, similar conditions of cell culture and treatment were determined suitable for the miRNA profiling experiments.

Known human miRNAs number in the thousands, so it was deemed sensible to begin with a wide-scale assay of SFN's effects on miRNA expression in each cell line, in order to highlight miRNAs to which attention is best drawn. As discussed in previous chapters, there are several techniques for wide-scale miRNA expression profiling, the miRNA library cloning and deep sequencing approach of which was determined to be the most suitable. Following the wide-scale profiling process, miRNAs apparently modulated by SFN and relevant to cancer (according to reports in the literature) were selected for further study.

Due to the sheer complexity of data and required normalisation processes, analyses of miRNA library cloning and deep sequencing data inevitably generate artefacts i.e. false indications of modulation. Therefore, the actual SFN-mediated modulation of any miRNA of interest was confirmed by an individually targeted miRNA assay, such as RT-qPCR or the Northern Blot.

4.2 Caco-2 and CCD-841 MicroRNA Library Deep Sequencing

As outlined in the Methods chapter, Caco-2 and CCD-841 cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (10 μ M) or DMSO (control) for 24 h; final DMSO concentrations were 0.05% (v/v). Total

RNA was extracted 8 or 24 h following treatment, from which miRNA libraries were constructed, that were subsequently subjected to deep sequencing. For gel images and other data obtained during the cloning process, see the Appendix. The deep sequencing data were quality checked, and then processed to remove adapter-adapter reads, as outlined in Methods.

Caco-2 library data quality assessment:

Table 6: The identity of each Caco-2 miRNA library, the proportion of reads of each remaining after the quality check and removal of adapter-adapter reads, and the subsequent proportion of reads mapping to the human genome. Libraries were constructed with three biological replicates per time point; 'A' libraries represent controls, 'B' 8 h SFN treatment, and 'C' 24 h SFN treatment.

Lib.	Time (h)	Rep.	Raw Read Number	Proportion After QC and Adapter Removal (%)	Matching Human Genome (%)	Reads Remaining After t/rRNA Removal
A1	0	1	19 985 254	93	82	13 539 159
A2		2	26 152 561	95	80	19 211 757
A3		3	16 176 855	96	80	12 200 936
B1	8	1	16 002 804	93	78	11 096 534
B2		2	14 603 739	92	80	10 258 639
B3		3	13 554 322	93	78	9 484 717
C1	24	1	23 029 947	95	79	16 136 765
C2		2	22 223 346	95	79	15 459 590
C3		3	26 035 997	93	80	17 208 841

CCD-841 library data quality assessment:

Table 7: The identity of each CCD-841 miRNA library, the proportion of reads of each remaining after the quality check and removal of adapter-adapter reads, and the subsequent proportion of reads mapping to the human genome. Libraries highlighted in red were excluded from further analysis because very low proportions of their reads were mappable to the human genome, potentially due to contamination of samples with non-human RNA during the construction of these libraries. Libraries were constructed with three biological replicates per time point; 'X' libraries represent controls, 'Y' 8 h SFN treatment, and 'Z' 24 h SFN treatment.

Lib.	Time (h)	Rep.	Raw Read Number	Proportion After QC and Adapter Removal (%)	Matching Human Genome (%)	Reads Remaining After t/rRNA Removal
X1	0	1	20 670 894	88	81	10 604 163
X2		2	12 051 340	71	78	4 169 711
X3		3	15 407 379	82	79	6 710 224
Y1	8	1	26 704 185	98	9	-
Y2		2	28 258 638	98	6	-
Y3		3	21 036 399	59	76	6 120 616
Z1	24	1	24 901 235	52	78	10 597 858
Z2		2	14 611 109	96	82	8 441 483
Z3		3	34 231 130	96	63	17 598 456

The data from the CCD-841 libraries Y1 and Y2 would not be meaningful for miRNA expression profiling, as indicated by several factors:

- The size class and complexity distributions do not at all resemble those expected for human miRNAs (see Appendix for distribution data and further detail).
- Very low proportions of reads are mappable to the human genome (see Table 7).

These issues may indicate heavy contamination with degraded RNA from another species during miRNA library preparation. Therefore, these two libraries were disregarded.

Inter-replicate library similarity was evaluated as described in Methods. Libraries A3 (Caco-2; 0 h) and B3 (Caco-2; 8 h) were hence excluded from further analysis, since they were significantly dissimilar to their replicate counterparts according to the Jaccard top 1000 and box plot tests. Library Z3 (CCD-841; 24 h) was also excluded, since it had more than double the number of genome-mapping reads vs. its replicate counterparts: Z1 and Z2.

Library data were then normalised, and SFN-mediated miRNA modulation at the two time points was determined as described in Methods. The names of all miRNAs thus determined as SFN-modulated in Caco-2 and CCD-841 cells are listed in Table 9 and Table 10 respectively, alongside the medians of their respective normalised read numbers across replicates.

Table 8: A summary of the numbers of total and differentially expressed known miRNAs detected in each cell line.

MicroRNA Library Set	Total MicroRNAs Detected	DE MicroRNAs	Proportion of MicroRNAs Apparently DE
Caco-2	1288	52	4.0%
CCD-841	1014	113	11.1%

4.2.1 MicroRNAs Modulated by Sulforaphane Treatment in Caco-2 According to Libraries

Table 9: The identities and normalised read number medians for miRNAs in Caco-2 for which the calculated fold change of expression between 0 and 8 h, 8 and 24 h, or 0 and 24 h of SFN (10µM) treatment was ≥ 1.9 . vA = Median((Reads in A1),(Reads in A2)); vB = Median((Reads in B1),(Reads in B2)); vC = Median((Reads in C1),(Reads in C2),(Reads in C3)).

MicroRNA	vA (0 h)	vB (8 h)	vC (24 h)	Total
hsa-let-7a-5p MIMAT0000062	57905	77092	113506	248503

(Table 9 Continued)

MicroRNA	vA (0 h)	vB (8 h)	vC (24 h)	Total
hsa-let-7b-5p MIMAT0000063	375	381	984	1740
hsa-let-7d-3p MIMAT0004484	742	503	331	1576
hsa-let-7e-5p MIMAT0000066	26519	37243	56000	119762
hsa-let-7f-5p MIMAT0000067	55615	75116	110493	241224
hsa-let-7g-5p MIMAT0000414	55595	75115	110355	241065
hsa-miR-9-5p MIMAT0000441	1760	2549	3391	7699
hsa-miR-10a-5p MIMAT0000253	31055	38797	63272	133124
hsa-miR-10b-5p MIMAT0000254	31054	38797	63272	133123
hsa-miR-15a-3p MIMAT0004488	46	27	10	83
hsa-miR-17-3p MIMAT0000071	5064	3256	2557	10877
hsa-miR-18a-3p MIMAT0002891	1986	1352	930	4268
hsa-miR-19b-1-5p MIMAT0004491	585	414	255	1253
hsa-miR-19b-2-5p MIMAT0004492	104	37	41	182
hsa-miR-20a-3p MIMAT0004493	333	251	119	703
hsa-miR-25-5p MIMAT0004498	657	272	302	1231
hsa-miR-27a-5p MIMAT0004501	407	130	248	785
hsa-miR-27b-5p MIMAT0004588	1898	580	1100	3577
hsa-miR-33b-3p MIMAT0004811	510	128	366	1003
hsa-miR-34a-3p MIMAT0004557	221	153	89	463
hsa-miR-98-5p MIMAT0000096	388	394	1022	1804
hsa-miR-181a-3p MIMAT0000270	2154	1253	780	4187
hsa-miR-188-3p MIMAT0004613	98	49	40	187
hsa-miR-193b-3p MIMAT0002819	7353	4880	3767	16000
hsa-miR-197-3p MIMAT0000227	138	104	47	289
hsa-miR-200c-5p MIMAT0004657	2206	1160	950	4315
hsa-miR-210-5p MIMAT0026475	110	43	34	187
hsa-miR-219a-1-3p MIMAT0004567	68	83	152	303
hsa-miR-296-5p MIMAT0000690	1466	758	654	2878
hsa-miR-331-3p MIMAT0000760	4493	2540	2039	9071
hsa-miR-339-5p MIMAT0000764	5688	3174	1914	10776
hsa-miR-504-5p MIMAT0002875	54	25	12	90
hsa-miR-532-3p MIMAT0004780	5377	3753	2648	11778
hsa-miR-548b-3p MIMAT0003254	201	133	95	429
hsa-miR-590-5p MIMAT0003258	248	172	118	538
hsa-miR-625-3p MIMAT0004808	2570	1516	836	4921
hsa-miR-934 MIMAT0004977	5793	3846	2637	12275
hsa-miR-1257 MIMAT0005908	242	196	450	888
hsa-miR-1296-5p MIMAT0005794	8151	4639	3792	16582
hsa-miR-1297 MIMAT0005886	55	41	154	250
hsa-miR-1827 MIMAT0006767	102	112	399	613
hsa-miR-2277-3p MIMAT0011777	95	57	32	184

(Table 9 Continued)

MicroRNA	vA (0 h)	vB (8 h)	vC (24 h)	Total
hsa-miR-3135b MIMAT0018985	264	118	76	458
hsa-miR-3200-3p MIMAT0015085	157	108	69	334
hsa-miR-3611 MIMAT0017988	38	22	10	69
hsa-miR-4286 MIMAT0016916	4671	1613	3172	9456
hsa-miR-4454 MIMAT0018976	154	50	99	303
hsa-miR-4517 MIMAT0019054	74	121	41	236
hsa-miR-5585-3p MIMAT0022286	47	15	45	107
hsa-miR-6130 MIMAT0024614	102	112	399	613
hsa-miR-7975 MIMAT0031178	144	42	84	270
hsa-miR-7977 MIMAT0031180	733	361	534	1628

4.2.2 MicroRNAs Modulated by Sulforaphane Treatment in CCD-841 According to Libraries

Table 10: The identities and normalised read number medians for miRNAs in CCD-841 for which the calculated fold change of expression between 0 and 8 h, 8 and 24 h, or 0 and 24 h of SFN (10µM) treatment was ≥ 1.9 . vX = Median((Reads in X1),(Reads in X2),(Reads in X3)); vY = (Reads in Y3); vZ = Median((Reads in Z1),(Reads in Z2)).

MicroRNA	vX (0 h)	vY (8 h)	vZ (24 h)	Total
hsa-let-7d-3p MIMAT0004484	1106	700	1679	3485
hsa-miR-9-5p MIMAT0000441	161	230	13	404
hsa-miR-9-3p MIMAT0000442	42	54	8	104
hsa-miR-10a-5p MIMAT0000253	188157	223910	83597	495664
hsa-miR-10b-5p MIMAT0000254	187938	223699	83531	495168
hsa-miR-17-3p MIMAT0000071	143	117	316	576
hsa-miR-18a-5p MIMAT0000072	549	455	1036	2040
hsa-miR-19a-3p MIMAT0000073	2446	2040	4437	8923
hsa-miR-19b-3p MIMAT0000074	2446	2040	4437	8923
hsa-miR-27a-5p MIMAT0004501	325	355	696	1376
hsa-miR-29a-5p MIMAT0004503	553	494	1001	2048
hsa-miR-29b-3p MIMAT0000100	30838	21374	48781	100993
hsa-miR-31-5p MIMAT0000089	13418	14309	6773	34500
hsa-miR-31-3p MIMAT0004504	1020	883	516	2419
hsa-miR-32-5p MIMAT0000090	417	275	858	1550
hsa-miR-33a-5p MIMAT0000091	125	42	103	270
hsa-miR-33b-5p MIMAT0003301	43	19	55	117
hsa-miR-33b-3p MIMAT0004811	30	9	42	81
hsa-miR-34a-3p MIMAT0004557	65	44	111	220
hsa-miR-34b-5p MIMAT0000685	20	20	236	276
hsa-miR-34b-3p MIMAT0004676	11	6	35	52
hsa-miR-34c-5p MIMAT0000686	90	65	731	886
hsa-miR-34c-3p MIMAT0004677	27	33	89	149
hsa-miR-96-5p MIMAT0000095	1048	1292	299	2639
hsa-miR-99a-5p MIMAT0000097	66887	70822	25579	163288
hsa-miR-100-5p MIMAT0000098	67082	71019	25643	163744
hsa-miR-101-3p MIMAT0000099	493	471	1036	2000
hsa-miR-103a-3p MIMAT0000101	16872	14658	29747	61277
hsa-miR-106b-5p MIMAT0000680	1878	1816	4934	8628
hsa-miR-126-5p MIMAT0000444	114	109	232	455
hsa-miR-130b-3p MIMAT0000691	4824	4776	11637	21237
hsa-miR-135a-5p MIMAT0000428	144	127	9	280
hsa-miR-135b-5p MIMAT0000758	144	127	9	280

(Table 10 Continued)

MicroRNA	vX (0 h)	vY (8 h)	vZ (24 h)	Total
hsa-miR-136-5p MIMAT0000448	1827	955	3817	6599
hsa-miR-136-3p MIMAT0004606	1519	1304	2531	5354
hsa-miR-138-5p MIMAT0000430	10001	8572	3744	22317
hsa-miR-138-1-3p MIMAT0004607	183	180	52	415
hsa-miR-143-5p MIMAT0004599	374	214	614	1202
hsa-miR-146a-5p MIMAT0000449	1455	1932	946	4333
hsa-miR-152-5p MIMAT0026479	26	16	55	97
hsa-miR-154-5p MIMAT0000452	541	384	965	1890
hsa-miR-155-5p MIMAT0000646	5434	6991	2059	14484
hsa-miR-181a-5p MIMAT0000256	42655	44144	83238	170037
hsa-miR-182-5p MIMAT0000259	13615	18594	2490	34699
hsa-miR-183-5p MIMAT0000261	1258	1214	278	2750
hsa-miR-183-3p MIMAT0004560	42	26	10	78
hsa-miR-193a-5p MIMAT0004614	638	570	194	1402
hsa-miR-193b-3p MIMAT0002819	2145	1449	2852	6446
hsa-miR-194-5p MIMAT0000460	188	311	459	958
hsa-miR-199a-5p MIMAT0000231	14195	16450	41602	72247
hsa-miR-199b-5p MIMAT0000263	14059	16257	41319	71635
hsa-miR-203a-3p MIMAT0000264	34	27	7	68
hsa-miR-210-3p MIMAT0000267	340	240	513	1093
hsa-miR-218-5p MIMAT0000275	8	10	145	163
hsa-miR-221-5p MIMAT0004568	266	266	105	637
hsa-miR-221-3p MIMAT0000278	127694	142812	72165	342671
hsa-miR-296-3p MIMAT0004679	69	25	65	159
hsa-miR-299-3p MIMAT0000687	123	97	221	441
hsa-miR-301a-3p MIMAT0000688	475	432	1850	2757
hsa-miR-301b-3p MIMAT0004958	442	418	1728	2588
hsa-miR-331-3p MIMAT0000760	677	622	1262	2561
hsa-miR-340-5p MIMAT0004692	183	178	518	879
hsa-miR-342-3p MIMAT0000753	373	405	757	1535
hsa-miR-345-5p MIMAT0000772	640	519	1613	2772
hsa-miR-374b-3p MIMAT0004956	124	106	260	490
hsa-miR-376a-5p MIMAT0003386	180	151	352	683
hsa-miR-376c-5p MIMAT0022861	28	22	62	112
hsa-miR-424-5p MIMAT0001341	3490	2666	15112	21268
hsa-miR-424-3p MIMAT0004749	306	229	662	1197
hsa-miR-450a-5p MIMAT0001545	327	283	1473	2083
hsa-miR-450a-1-3p MIMAT0022700	13	14	61	88
hsa-miR-450a-2-3p MIMAT0031074	9	7	51	67
hsa-miR-450b-5p MIMAT0004909	142	140	634	916
hsa-miR-486-5p MIMAT0002177	18553	21591	3503	43647
hsa-miR-486-3p MIMAT0004762	603	761	202	1566
hsa-miR-490-3p MIMAT0002806	2267	1465	5152	8884

(Table 10 Continued)

MicroRNA	vX (0 h)	vY (8 h)	vZ (24 h)	Total
hsa-miR-497-5p MIMAT0002820	141	104	65	310
hsa-miR-500a-3p MIMAT0002871	1181	1177	465	2823
hsa-miR-503-5p MIMAT0002874	432	346	1793	2571
hsa-miR-504-5p MIMAT0002875	25	22	212	259
hsa-miR-542-5p MIMAT0003340	111	85	646	842
hsa-miR-542-3p MIMAT0003389	195	146	1068	1409
hsa-miR-543 MIMAT0004954	2787	1957	1306	6050
hsa-miR-550a-3p MIMAT0003257	143	92	228	463
hsa-miR-550b-2-5p MIMAT0022737	131	85	214	430
hsa-miR-561-5p MIMAT0022706	13	13	86	112
hsa-miR-584-5p MIMAT0003249	2473	2460	935	5868
hsa-miR-584-3p MIMAT0022708	38	32	10	80
hsa-miR-590-5p MIMAT0003258	23	7	66	96
hsa-miR-545-3p MIMAT0003165	11	1	36	48
hsa-miR-615-3p MIMAT0003283	1243	737	2023	4003
hsa-miR-618 MIMAT0003287	4	6	117	127
hsa-miR-625-5p MIMAT0003294	44	39	133	216
hsa-miR-671-5p MIMAT0003880	138	100	226	464
hsa-miR-708-5p MIMAT0004926	1	2	26	29
hsa-miR-744-5p MIMAT0004945	1662	1104	3505	6271
hsa-miR-769-5p MIMAT0003886	335	235	867	1437
hsa-miR-769-3p MIMAT0003887	45	40	135	220
hsa-miR-887-3p MIMAT0004951	49	46	199	294
hsa-miR-935 MIMAT0004978	63	41	22	126
hsa-miR-1185-1-3p MIMAT0022838	405	307	609	1321
hsa-miR-1185-2-3p MIMAT0022713	405	307	609	1321
hsa-miR-1197 MIMAT0005955	27	22	78	127
hsa-miR-1287-5p MIMAT0005878	50	50	141	241
hsa-miR-1296-5p MIMAT0005794	598	337	955	1890
hsa-miR-1301-3p MIMAT0005797	33	40	110	183
hsa-miR-1307-5p MIMAT0022727	121	106	295	522
hsa-miR-2110 MIMAT0010133	147	155	68	370
hsa-miR-2355-5p MIMAT0016895	42	38	97	177
hsa-miR-3074-5p MIMAT0019208	6361	9018	17243	32622
hsa-miR-3120-3p MIMAT0014982	14	16	56	86
hsa-miR-3934-5p MIMAT0018349	34	34	8	76
hsa-miR-6720-5p MIMAT0027345	11	2	39	52

4.2.3 Examination of the Library Data Analysis and IsomiRs

For each miRNA named in miRBase (e.g. hsa-miR-29b-3p), only the sequence of the canonical form is presently catalogued. However, each

miRNA has multiple isoforms (isomiRs), whose sequences vary from those of the canonical forms by having extra or missing nucleotides at the 5' and 3' ends, or substituted bases in some cases. These originate, in part, from the 'imprecise' cleavage of pri-miRNAs and pre-miRNAs by Drosha and Dicer respectively, generating end deletions or templated additions. Another source of variation is the susceptibility of mature miRNAs and pre-miRNAs to post-transcriptional modification by exonucleases and/or base-modifying enzymes. The canonical miRNA sequence catalogued in miRBase tends to be either a consensus from pooled sequencing data, or the form most abundant in the experimental data from which the miRNA acceded to miRBase.

The algorithm that was used to match sequencing reads to miRBase-catalogued human miRNAs allowed up to 2 mismatches, in order to ensure the inclusion of non-canonical isomiRs. As a result, the summed read numbers reported for "hsa-miR-29b-3p" – for example – were not all from the miRBase-canonical isoform of hsa-miR-29b-3p, but included reads from various isomiRs. Non-canonical isomiRs are important to consider, because they can be much more abundant than their canonical counterparts in some contexts.

Unfortunately, some distinct miRNAs differ from one another in canonical sequence by only 1 or 2 nucleotides; the tolerance of mismatches inevitably leads to some misannotation. The annotated library data first appeared to imply that hsa-let-7a-5p was highly abundant and upregulated by SFN. However, further inspection of the reads annotated as hsa-let-7a-5p revealed that none of them actually matched the hsa-let-7a-5p sequence. Most of the reads strictly matched the canonical hsa-let-7f-5p and hsa-let-7g-5p sequences, each of which differ in sequence from hsa-let-7a-5p by only 1 and 2 nucleotides respectively, as illustrated in Table 11.

Table 11: The canonical sequences of hsa-let-7a-5p, hsa-let-7f-5p and hsa-let-7g-5p.

7a	u	g	a	g	g	u	a	g	u	a	g	g	u	u	g	u	a	u	a	g	u	u
7f	u	g	a	g	g	u	a	g	u	a	g	a	u	u	g	u	a	u	a	g	u	u
7g	u	g	a	g	g	u	a	g	u	a	g	u	u	u	g	u	a	c	a	g	u	u

Therefore, the summed reads for miRNAs of interest were carefully scrutinised before further study. Incorrect annotation was relatively rare, and in most cases, the majority (70-80%) of the reads annotated as a specific miRBase entry were in fact perfect matches to the canonical sequence. However non-canonical isomiRs dominated in some contexts; these isomiRs varied from the canonicals by having 1 or 2 extra or missing nucleotides at the (often 3') end.

For example, of the Caco-2 library reads annotated as hsa-miR-17-3p:

- 43% represented an isomiR with an extra nucleotide at the 3' end.
- 18% perfectly matched the canonical sequence.
- 18% represented an isomiR missing a nucleotide at the 3' end.
- 10% represented an isomiR missing 2 nucleotides at the 3' end.
- 5% represented an isomiR missing one nucleotide at the 5' end.

Thus, it was decided that the first isomiR of miR-17-3p would be more appropriate for validation due to its apparent dominance. Using the isomiR nomenclature proposed by Coonan et al. and described in MicroRNA Nomenclature in the Appendix, this isomiR was named hsa-miR-17-3p[[isomiR]]50_72|.

4.3 Validating Differential MicroRNA Expression in Response to Sulforaphane

Northern Blotting and TaqMan RT-qPCR were used to assay SFN's effects on miRNA expression individually. Caco-2 and CCD-841 cells were treated with SFN or DMSO (for controls); final DMSO concentrations were 0.05% (v/v). Northern Blot experiments were performed in duplicate, whilst RT-

qPCR experiments were performed in triplicate. U6 was used as an internal control for normalisation.

4.3.1 The Let-7 Family

Anti-cancer effects of the let-7 miRNAs are frequently reported, due to the tendency of these miRNAs to repress proliferation, promote differentiation and/or promote apoptosis, particularly in CRC cells. Reads perfectly matching the canonical sequences of hsa-let-7f-5p and hsa-let-7g-5p were highly abundant in the data from both cell lines.

4.3.1.1 Let-7f-5p in Caco-2

The Caco-2 library data suggested that SFN (10 μ M) treatment induced let-7f-5p 1.4- and 2.1-fold at 8 and 24 h respectively, as illustrated in Figure 33.

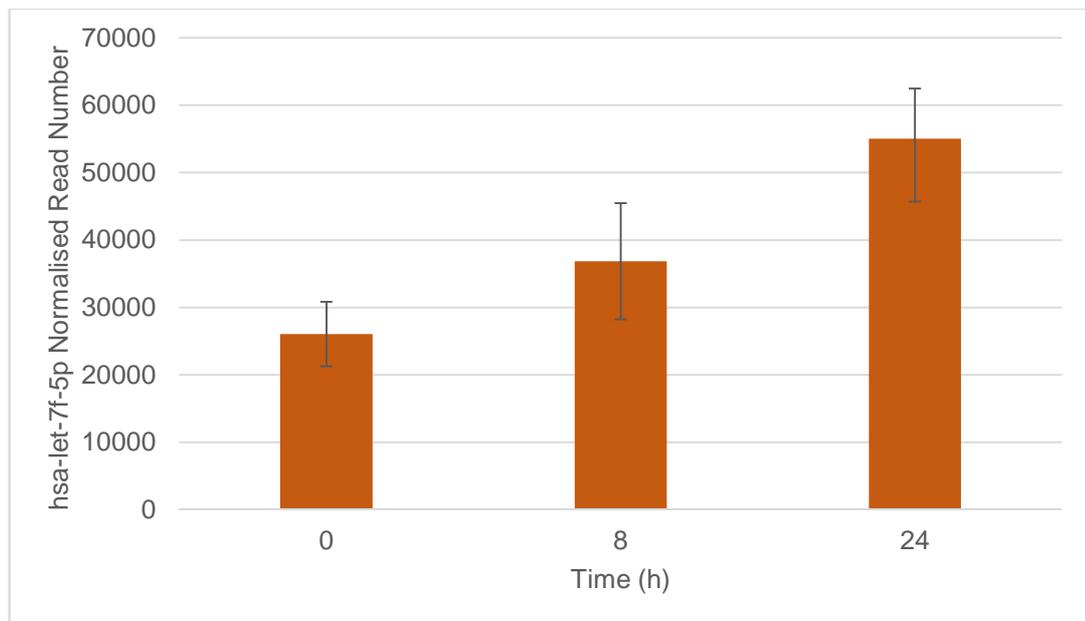


Figure 33: Normalised read numbers for hsa-let-7f-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

The Northern Blot data for let-7f-5p, as illustrated in Figure 34 corroborate a sequential trend of upregulation, albeit with more subtle inductions of 1.2- and 1.4-fold at 8 and 24 h respectively.

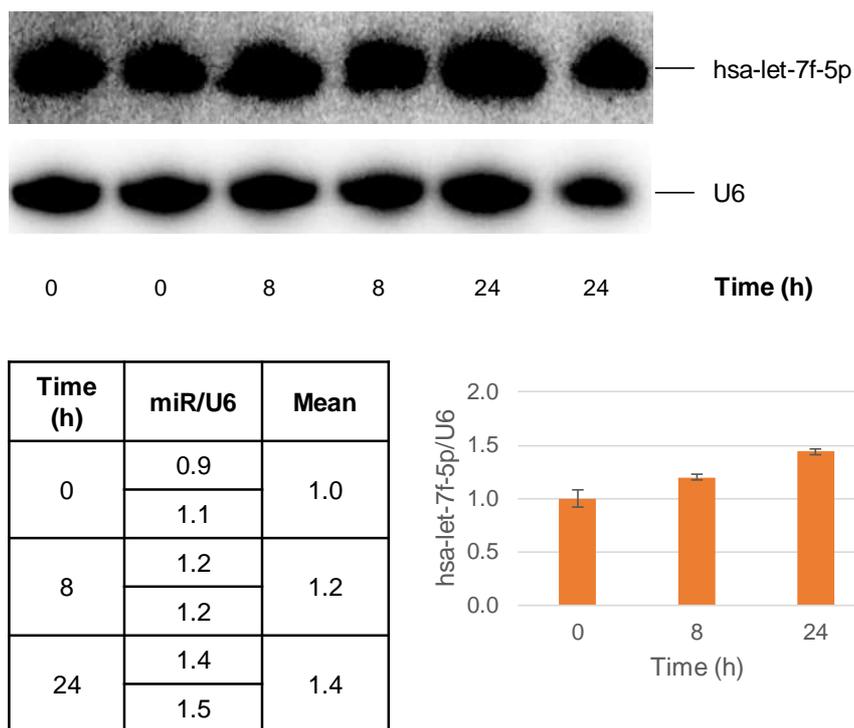


Figure 34: The effect of SFN treatment on the level of hsa-let-7f-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.

4.3.1.2 Let-7g-5p in Caco-2

The Caco-2 library data suggested that let-7g-5p was also upregulated by SFN (10 μ M) treatment, 1.3- and 2.0-fold at 8 and 24 h respectively, as illustrated in Figure 35.

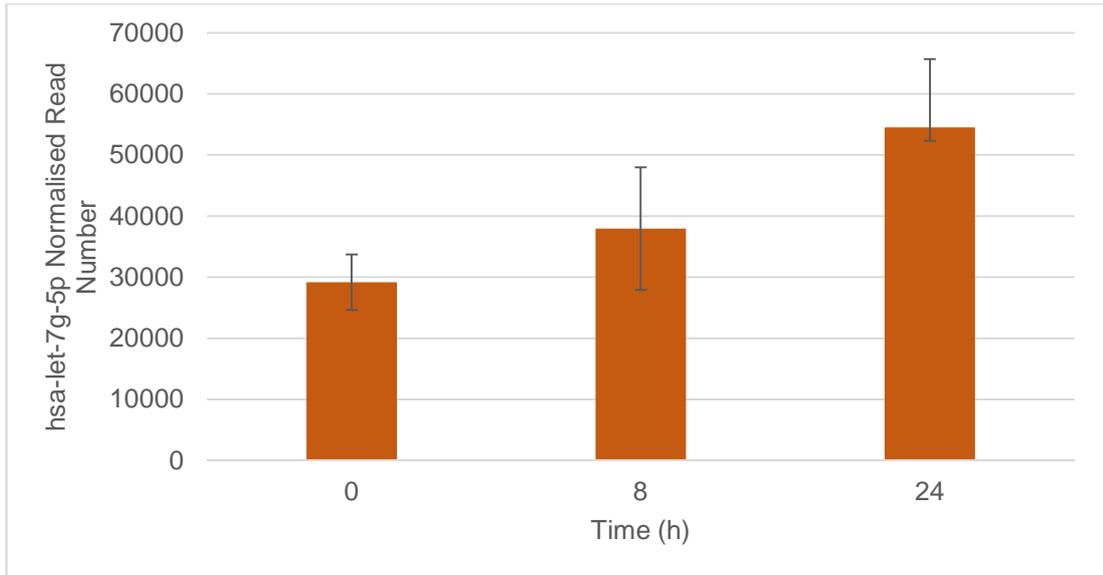


Figure 35: Normalised read numbers for hsa-let-7g-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

The Northern Blot data for let-7g-5p as illustrated in Figure 36 imply inductions of 1.1- and 1.6-fold at 8 and 24 h respectively, thus confirming that let-7g-5p expression is upregulated by SFN treatment at 24 h.

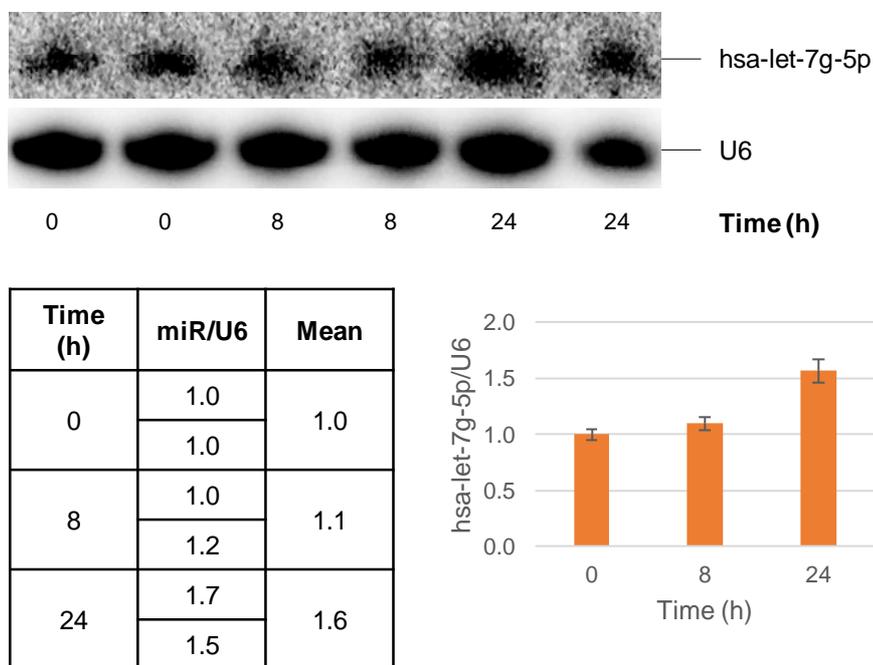


Figure 36: The effect of SFN treatment on the level of hsa-let-7g-5p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.

4.3.1.3 Let-7f-5p and Let-7g-5p in CCD-841

The effects of SFN treatment on the expression of let-7f-5p and let-7g-5p in CCD-841 was subsequently tested, to determine whether or not they responded similarly to the Caco-2 cells. As illustrated in Figure 39 and Figure 40, the Northern Blot data actually indicate that SFN does not affect these two miRNAs in CCD-841, as corroborated by the CCD-841 library data read numbers (which are illustrated in Figure 37 and Figure 38).

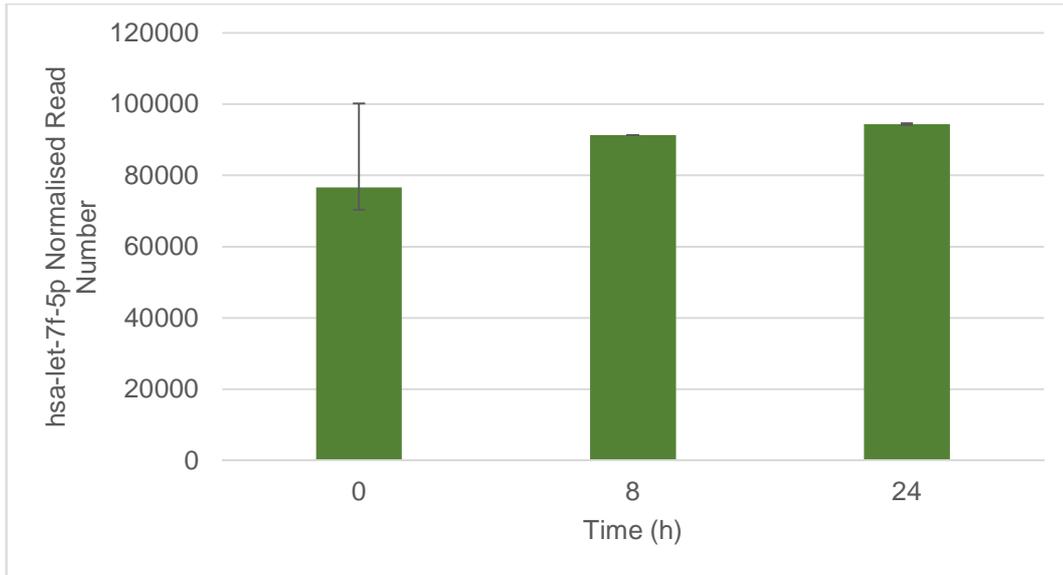


Figure 37: Normalised read numbers for hsa-let-7f-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

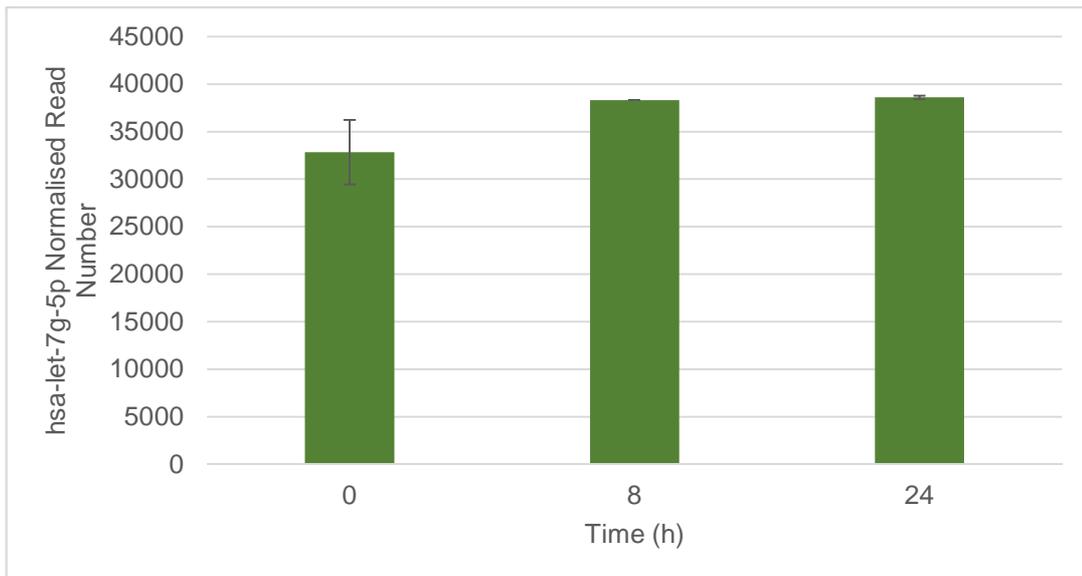
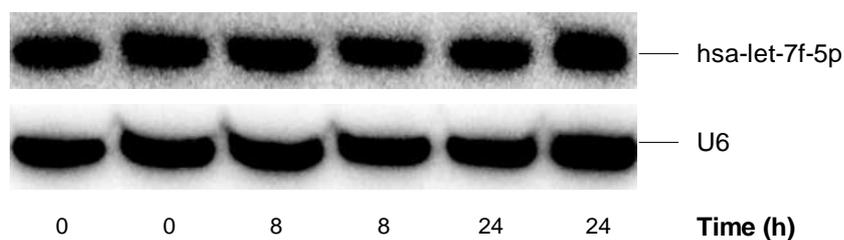


Figure 38: Normalised read numbers for hsa-let-7g-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

Let-7f-5p and let-7g-5p have important anti-proliferative, pro-differentiative and pro-apoptotic functions, so it is potentially significant that their induction by SFN appears to occur selectively in the CRC and not the non-cancerous

colonic cell line. This could have implications for SFN's therapeutic potential.



Time (h)	miR/U6	Mean
0	1.0	1.0
	1.0	
8	1.0	1.0
	0.9	
24	1.0	1.0
	1.0	

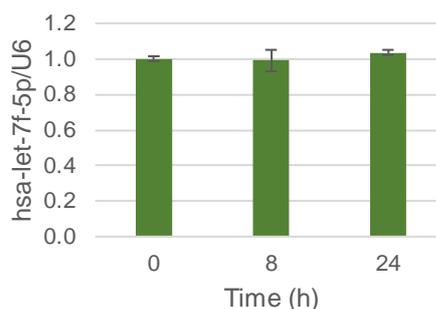
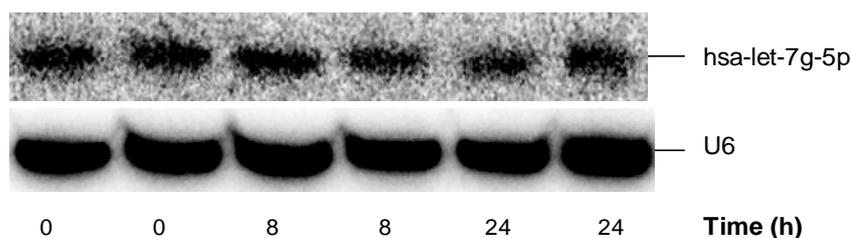


Figure 39: The effect of SFN treatment on the level of hsa-let-7f-5p in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.



Time (h)	miR/U6	Mean
0	1.1	1.0
	0.9	
8	1.2	1.0
	0.8	
24	1.0	0.9
	0.8	

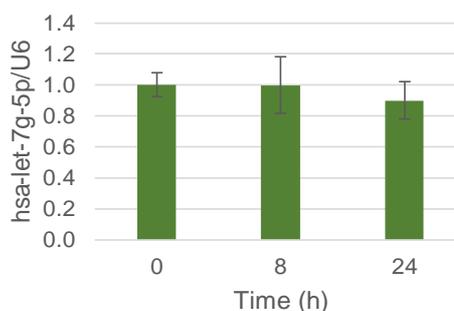
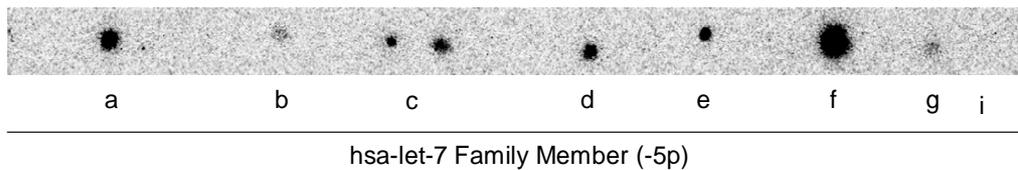


Figure 40: The effect of SFN treatment on the level of hsa-let-7g-5p in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.

It was speculated that the Northern Blot data may have ‘underreported’ the fold inductions of let-7f-5p and let-7g-5p, if the anti-let-7f-5p and anti-let-7g-5p probes cross-hybridised with other miRNAs that were close in sequence. Particularly similar in sequence are other members of the let-7 family, including hsa-let-7x-5p where x = a, b, c, d, e, f, g, i.

An approach for testing the potential for such cross-hybridisation is to make dot-blot by blotting positive-sense oligomers composed of the sequences of let-7x-5p, where x = a, b, c, d, e, f, g, i. The dot-blot are then probed using the anti-let-7f-5p and anti-let-7g-5p Northern Blot probes to compare the intensity of binding between the different blotted oligomers.

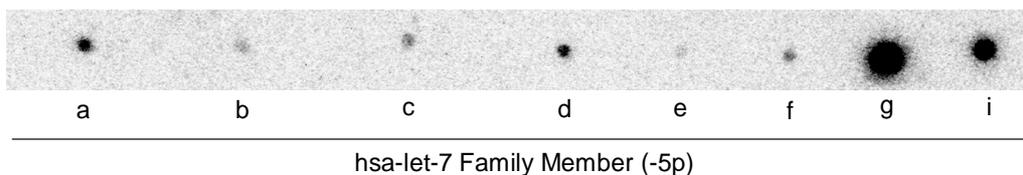
4.3.1.4 Dot-blot Probed with Anti-Let-7f-5p Probe



Let-7 Member	a	b	c	d	e	f	g	i
Relative Density (%)	16	1	3	3	3	100	1	0

Figure 41: A test of the cross-reactivity of the probe anti-sense to hsa-let-7f-5p with other let-7 family members. Positive sense oligomers composed of the sequences of each miRNA were blotted onto a membrane, which was then probed against hsa-let-7f-5p.

4.3.1.5 Dot blots Probed with Anti-Let-7g-5p Probe



Let-7 Member	a	b	c	d	e	f	g	i
Relative Density (%)	2	1	0	1	0	0	100	19

Figure 42: A test of the cross-reactivity of the probe anti-sense to hsa-let-7g-5p with other let-7 family members. Positive sense oligomers composed of the sequences of each miRNA were blotted onto a membrane, which was then probed against hsa-let-7g-5p.

The dot-blot data illustrated in Figure 41 and Figure 42 indicate little cross-hybridisation between the anti-let-7f-5p and anti-let-7g-5p probes, and other let-7 miRNA dot-blot. Notably, almost no reactivity between anti-let-7f-5p and let-7g-5p, or between anti-let-7g-5p and let-7f-5p, is observable. This suggests that both miRNAs are really upregulated, rather than the apparent upregulation of one being an artefact arising from cross-hybridisation of the probe targeted against the other. Both probes cross-reacted to a small extent with other let-7 members; the anti-let-7f-5p probe with let-7a-5p, and the anti-let-7g-5p probe with let-7i-5p. However, reads matching neither let-7a-5p nor

let-7i-5p were found in the Caco-2 library data, implying that these miRNAs were either absent or very low in abundance, and thus unlikely to confound the Northern Blot data.

4.3.2 The MiR-10 Family

Caco-2 library data reads annotated to miR-10a-5p and miR-10b-5p were abundant, and also implied that these miRNAs are regulated by SFN. Reports in the literature regarding the roles of miR-10a-5p and miR-10b-5p in cancer vary from oncogenic to tumour-suppressive.

Scrutiny of the miR-10a-5p-annotated reads reveals that those strictly matching the canonical sequence indicate induction by SFN (10 μ M) treatment of 1.1- and 1.5-fold at 8 and 24 h respectively, as illustrated in Figure 43.

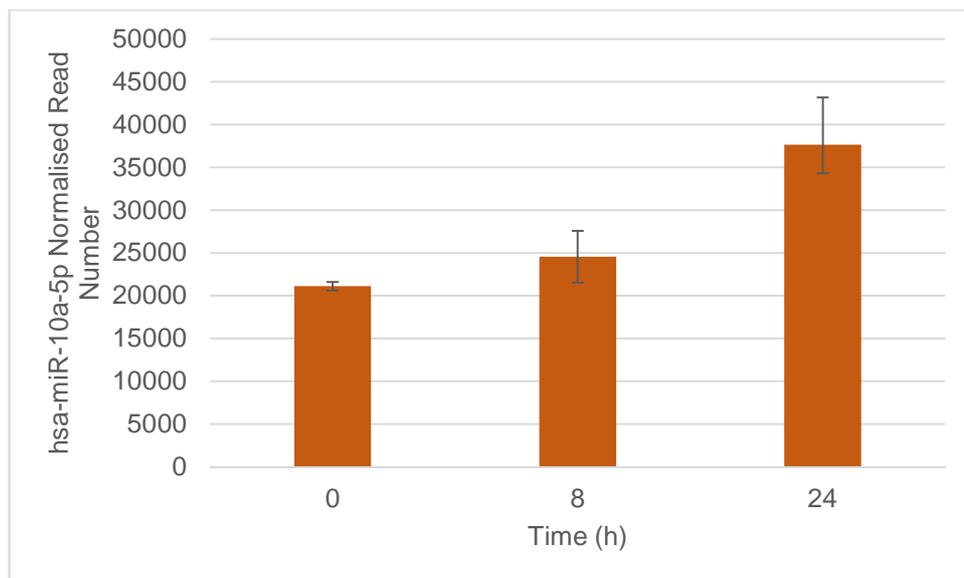


Figure 43: Normalised read numbers for hsa-miR-10a-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

The reads annotated as miR-10a-5p[[isomiR]|21_42] imply induction of 1.2- and 1.5-fold at 8 and 24 h respectively.

4.3.2.1 MiR-10a-5p in Caco-2

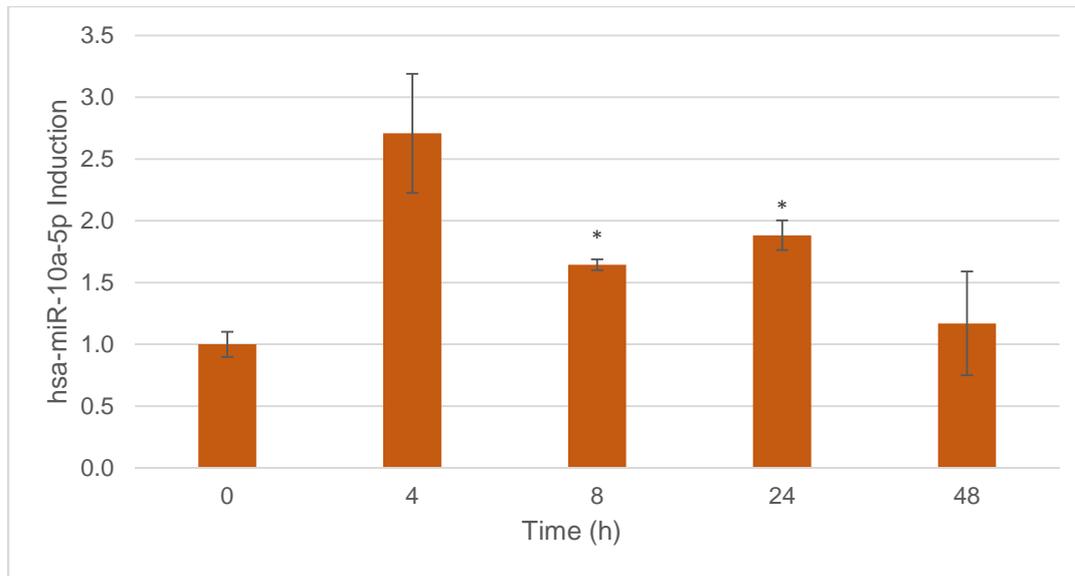


Figure 44: The effect of SFN treatment on the level of hsa-miR-10a-5p in Caco-2 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control for normalisation. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (* p <0.05).

The time-course RT-qPCR data illustrated in Figure 44 indicate a statistically significant (according to an unpaired two-tailed Student's T-test) induction of canonical miR-10a-5p in Caco-2, by 1.6- and 1.9-fold at the 8 and 24 h time points respectively, thus confirming that it is induced by SFN treatment. The trend observed in the data shown in Figure 45 suggests that there is a dose-dependent induction of miR-10a-5p by SFN (5 to 20 μ M) in Caco-2.

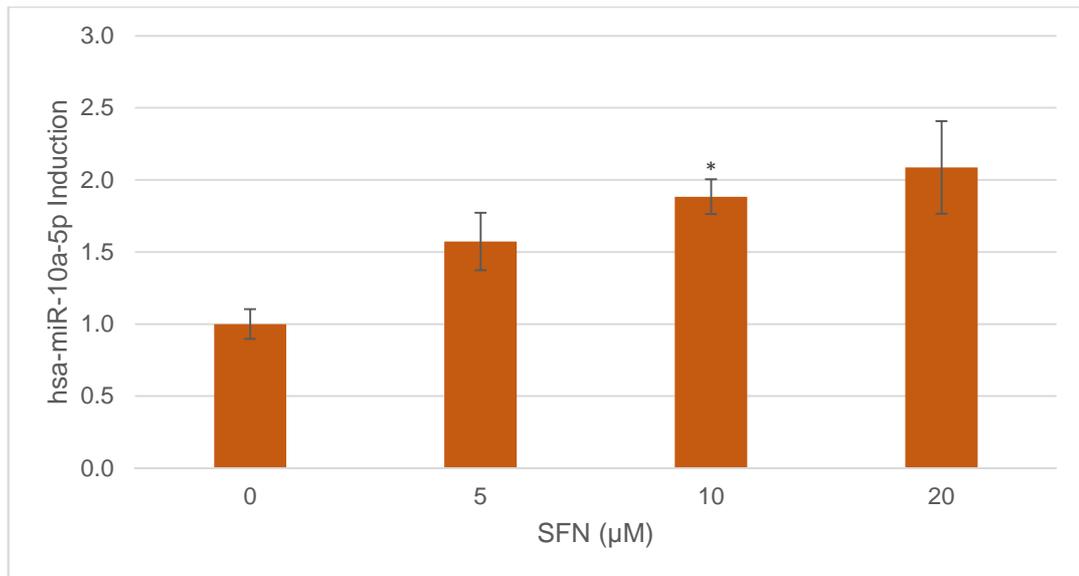


Figure 45: The effect of 24 h SFN treatment on the level of hsa-miR-10a-5p in Caco-2 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20μM) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control for normalisation. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates ± S.E.M. (*p<0.05).

The Northern Blot data illustrated in Figure 46 for miR-10a-5p partially corroborate its modulation by SFN at 8 h. However, in contrast to the Caco-2 library data, the Northern Blot data imply that expression at 24 h falls back to levels similar to those in controls.

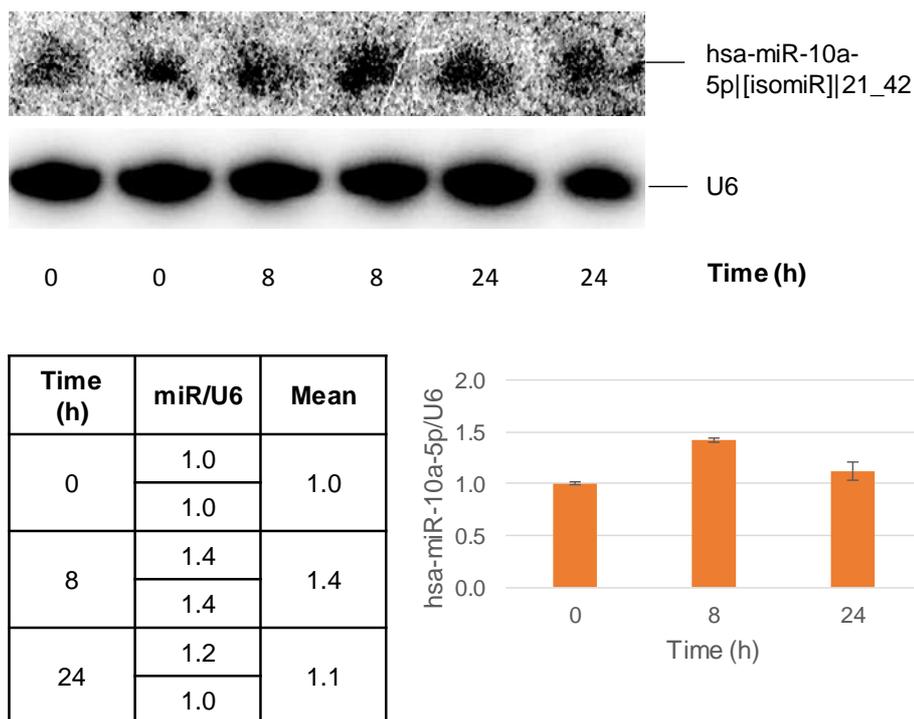


Figure 46: The effect of SFN treatment on the level of hsa-miR-10a-5p[[isomiR]]21_42 in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.

This also contrasts with the pattern of canonical miR-10a-5p expression observed in the RT-qPCR data illustrated in Figure 44. Perhaps extended SFN treatment promotes the dominance of one isomiR over the other. However, it is difficult to draw conclusions from the Northern Blot data shown in Figure 46 due to the weak band signals.

4.3.2.2 MiR-10b-5p in Caco-2

The Caco-2 library data indicated that canonical miR-10b-5p was induced by SFN 1.4- and 2.4-fold at 8 and 24 h respectively, as illustrated in Figure 47.

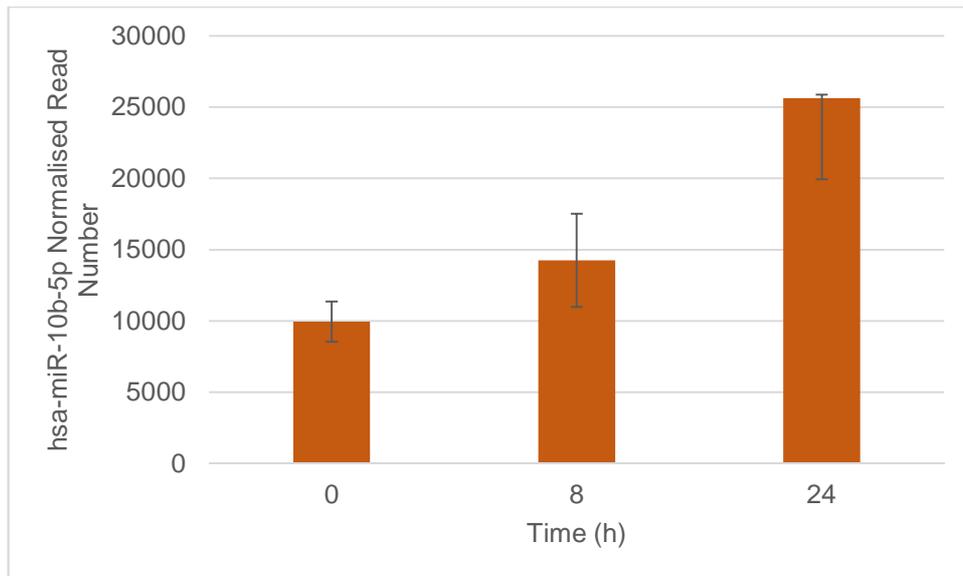


Figure 47: Normalised read numbers for hsa-miR-10b-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, the RT-qPCR data (see Appendix) did not show any significant differences nor trend. Likewise, the Northern Blot data for miR-10b-5p|[isomiR]|26_47| (see Appendix) did not corroborate the changes in expression of this miRNA inferred from the library data analysis.

4.3.3 MiR-29b-3p

MiR-29b-3p is an miRNA whose function apparently varies according to context, since it is reported as both a tumour suppressor and an oncogene. Interestingly, it has been demonstrated to be repressed by Nrf2, probably at the transcriptional level via the binding of Nrf2 to an *ARE* in the promoter region of the miR-29b-1 gene (Shah, Zaitseva et al. 2015). The Northern Blot data illustrated in Figure 48 show that this miRNA is downregulated by SFN treatment in Caco-2.

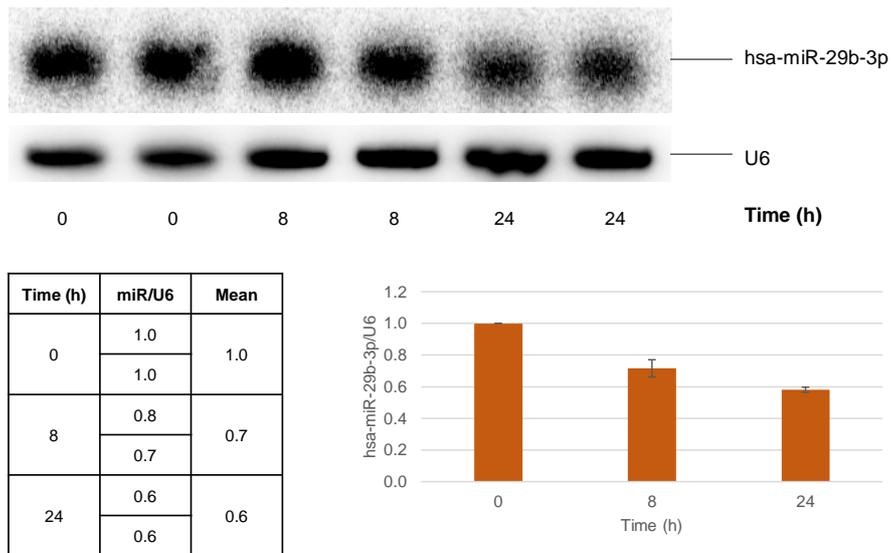


Figure 48: The effect of SFN treatment on the level of hsa-miR-29b-3p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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.

Analysis of the Caco-2 library data did not actually imply the modulation of miR-29b-3p by SFN, although such modulation may have been masked by high inter-replicate read number variation, as illustrated in Figure 49.

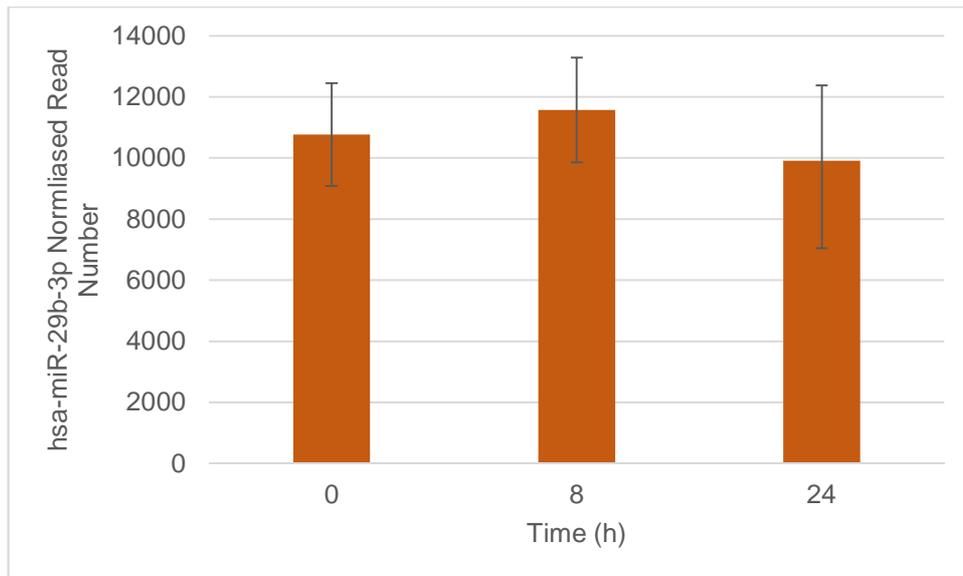


Figure 49: Normalised read numbers for hsa-miR-29b-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

The implications of the finding that SFN treatment downregulates miR-29b-3p expression in Caco-2 cells are unclear, since this miRNA has received both tumour suppressive and oncogenic reports. For example, it has been demonstrated to suppress EMT in CRC by targeting Tiam1 (Wang, Li et al. 2014), and to inhibit the Wnt signalling pathway (Subramanian, Rao et al. 2014). On the other hand, it has been reported to be overexpressed in, and enhance the invasiveness of, breast cancer cells (Yan, Guo et al. 2015).

Interestingly, a different ITC – PEITC – has been shown to inhibit the environmental cigarette smoke-mediated upregulation of miR-29b-3p in the mouse lung (Izzotti, Larghero et al. 2010). The ability of ITCs such as SFN and PEITC to downregulate miR-29b-3p may be due to Nrf2 induction, since Nrf2 has been demonstrated to be able to downregulate miR-29b-3p by repressing transcription at the miR-29b-1 genetic locus in acute myeloid leukaemia cells (Shah, Zaitseva et al. 2015).

4.3.4 MiR-193b-3p

MiR-193b-3p is a potential oncomiR that was indicated by the library data to be downregulated by SFN (10 μ M) treatment 1.5- and 2.0-fold at 8 and 24 h respectively, as illustrated in Figure 50.

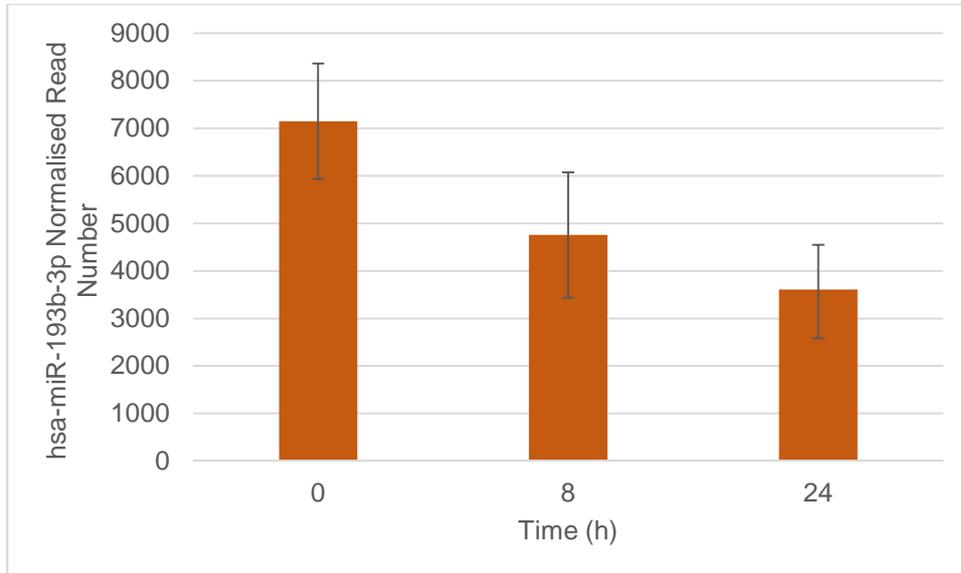


Figure 50: Normalised read numbers for hsa-miR-193b-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, the Northern Blot data as illustrated in Figure 51 do not strongly support this; they corroborate downregulation at 8 h, but not at 24 h. This miRNA is similar in sequence to hsa-miR-193a-3p. However, the lack of effect observable at 24 h in the Northern Blot is unlikely to be an artefact of the anti-miR-193b-3p probe cross-hybridising with miR-193a-3p, because the library data suggested that the latter was also downregulated at 24 h.

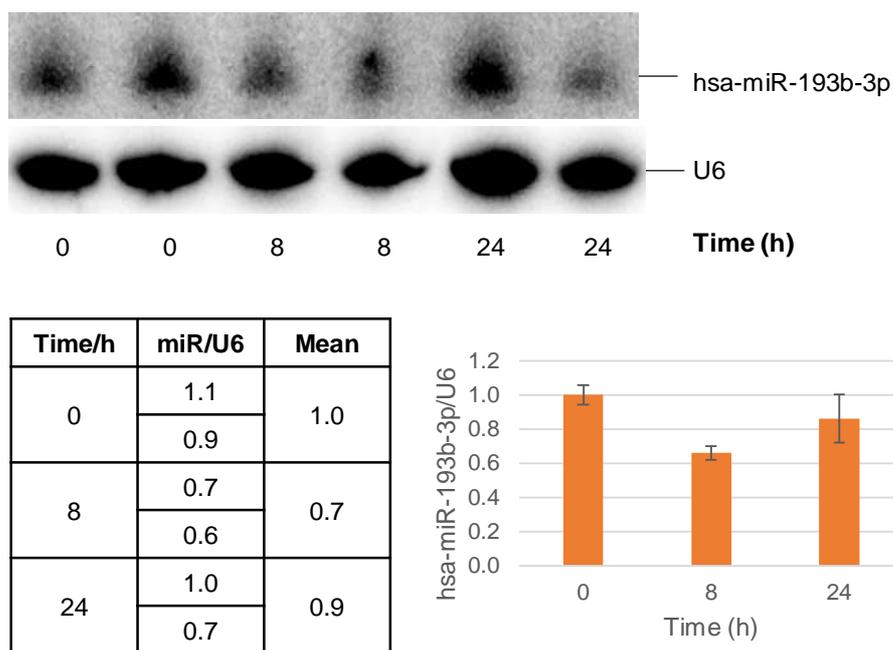


Figure 51: The effect of SFN treatment on the level of hsa-miR-193b-3p in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10µM) for 8 or 24 h, or DMSO; final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance was assayed via Northern Blotting as described in Methods, using U6 as an internal control for normalisation. 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 ± (upper value – lower value)/2.

As a potential oncomiR downregulated by SFN treatment, it would be interesting to validate the modulation of miR-193b-3p by additional Northern Blots and/or RT-qPCR experiments in the future, and then if it is confirmed to be differentially expressed, to study it further.

4.3.5 Other MiRNAs

4.3.5.1 Caco-2

The library data implied that miR-182-5p was upregulated by SFN (10µM) in Caco-2 by 1.4- and 1.6-fold at 8 and 24 h respectively, as illustrated in Figure 52.

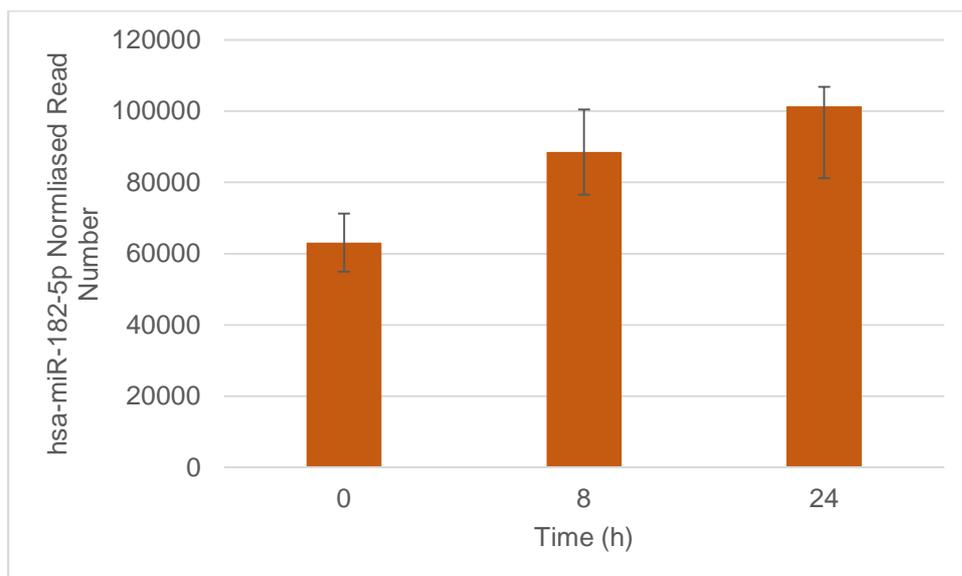


Figure 52: Normalised read numbers for hsa-miR-182-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, the Northern Blot (see Appendix for data) did not convincingly corroborate this. MiR-192-5p was also indicated by the library data as upregulated by SFN (10 μ M) 1.7-fold at 24 h, as illustrated in Figure 53, but this also was not supported by the Northern Blot (see Appendix for data).

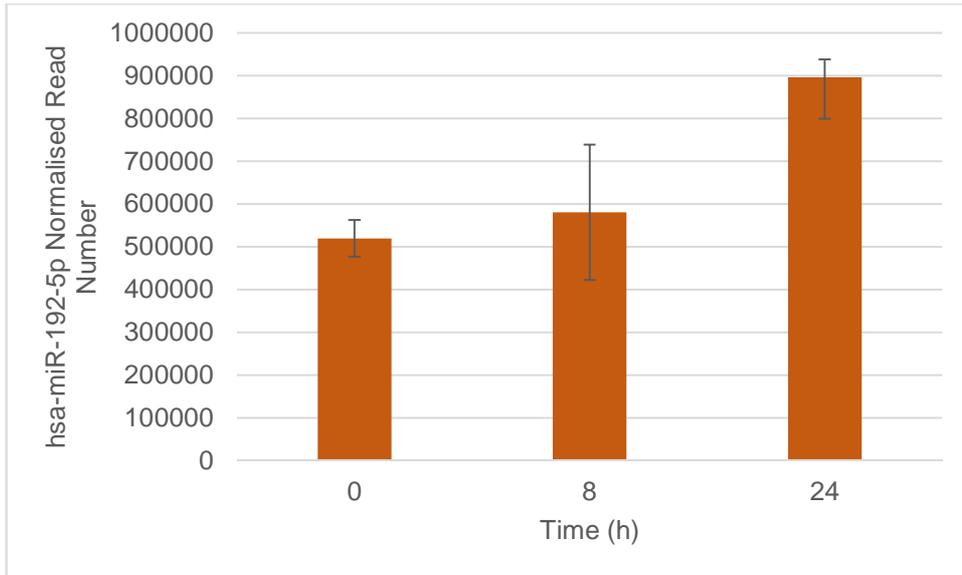


Figure 53: Normalised read numbers for hsa-miR-192-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

The library data had implied miR-17-3p to be downregulated by SFN (10 μ M) 1.6- and 2.0-fold at 8 and 24 h respectively, as illustrated in Figure 54; this neither was supported by the Northern Blot (see Appendix for data).

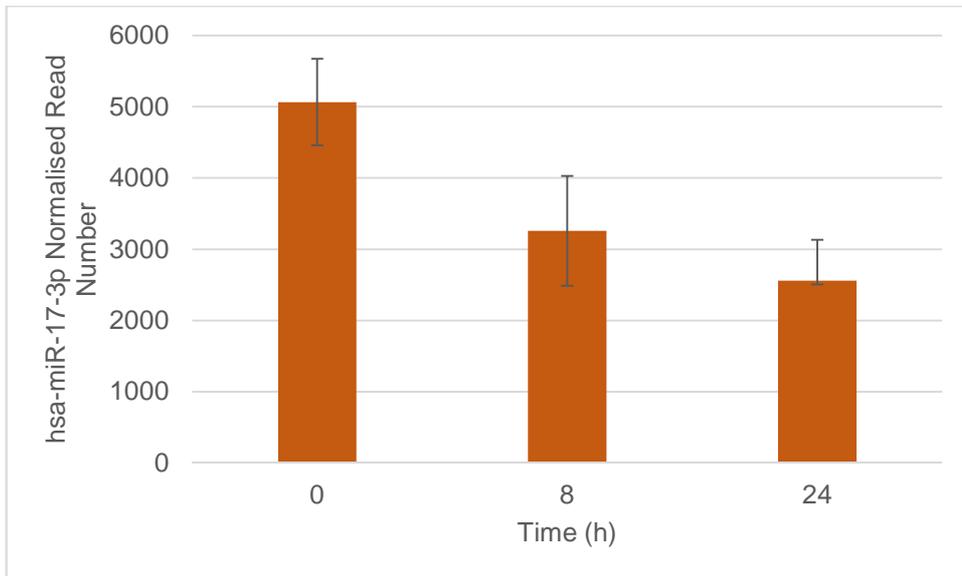


Figure 54: Normalised read numbers for hsa-miR-17-3p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

Finally, miR-1296-5p appeared to be downregulated by SFN as illustrated in Figure 55. The Northern Blot data (see Appendix) were unable to confirm this, although this may have been due to very weak band signals, which themselves imply low overall expression.

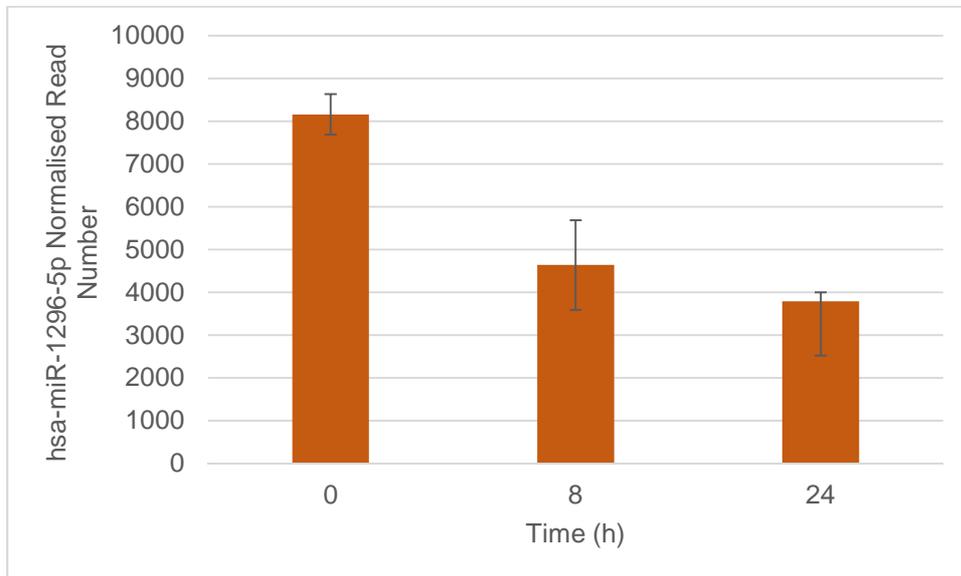


Figure 55: Normalised read numbers for hsa-miR-1296-5p in Caco-2 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

4.3.5.2 CCD-841

The library data indicated that hsa-miR-182-5p was upregulated in CCD-841 by SFN (10 μ M) 1.4-fold at 8 h, then downregulated 5.5-fold at 24 h, as illustrated in Figure 56.

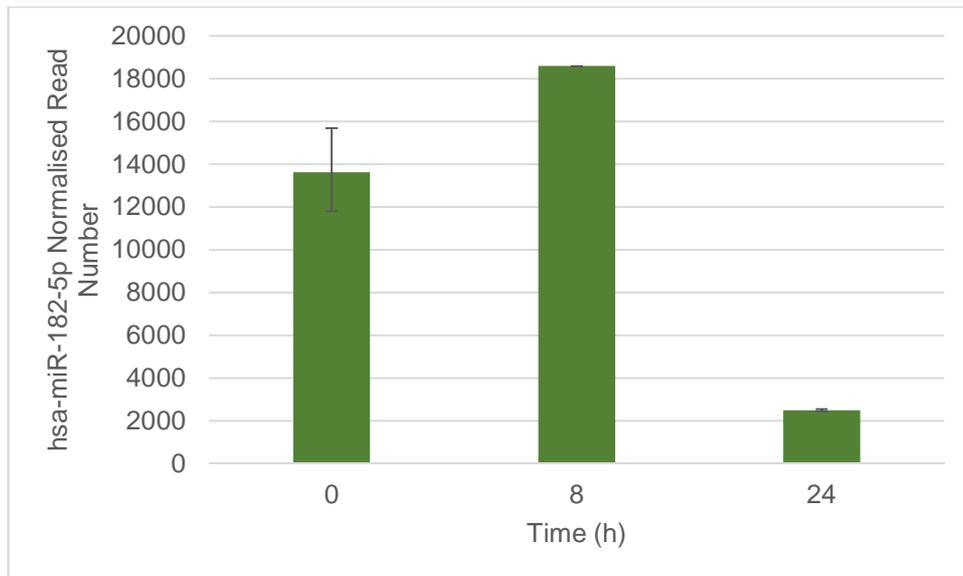


Figure 56: Normalised read numbers for hsa-miR-182-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, the Northern Blot (see Appendix for data) did not corroborate this. Although it is possible that the anti-miR-182-5p probe cross-hybridised with miR-183-5p due to sequence similarity, the latter was suggested to be downregulated in a similar fashion by the library data.

It appeared from library data analysis that miR-106b-5p was upregulated by SFN (10 μ M) in CCD-841 2.6-fold at 24 h (illustrated in Figure 57).

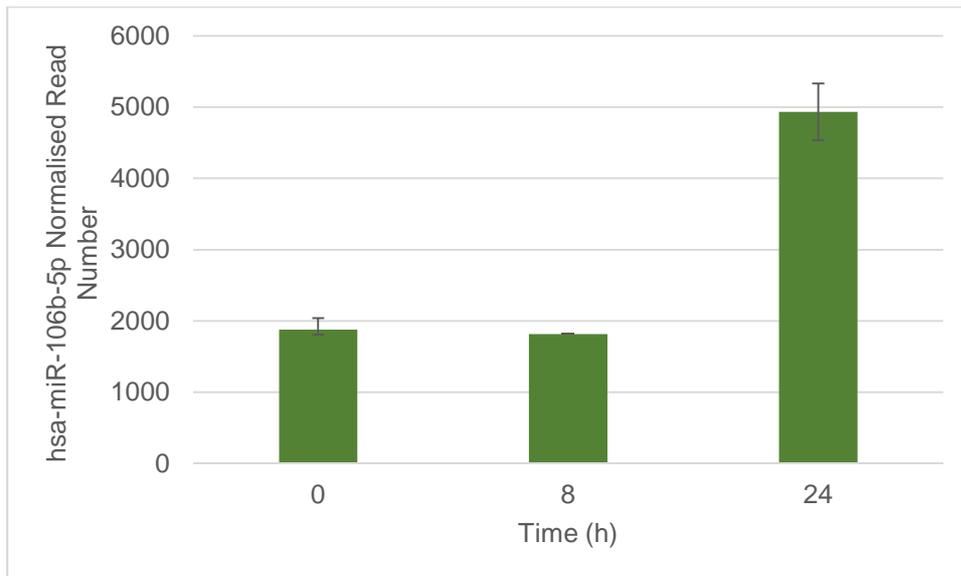


Figure 57: Normalised read numbers for hsa-miR-106b-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, this effect was not supported by the Northern Blot experiments (see Appendix for data). It is not clear whether or not this is an artefact of cross-reactivity of the anti-miR-106b-5p oligo with other close-in-sequence miRNAs such as miR-17-5p and miR-106a-5p, both of which have more reads in the library data and are not shown to be differentially expressed.

The upregulation of miR-181a-5p by SFN at 24 h in CCD-841 was inferred from the library data (as illustrated in Figure 58), but not confirmed by Northern Blot experiments. Nor did the time course Northern Blots indicate any changes at the 4, 8, 12 or 48 h time points (see Appendix for data).

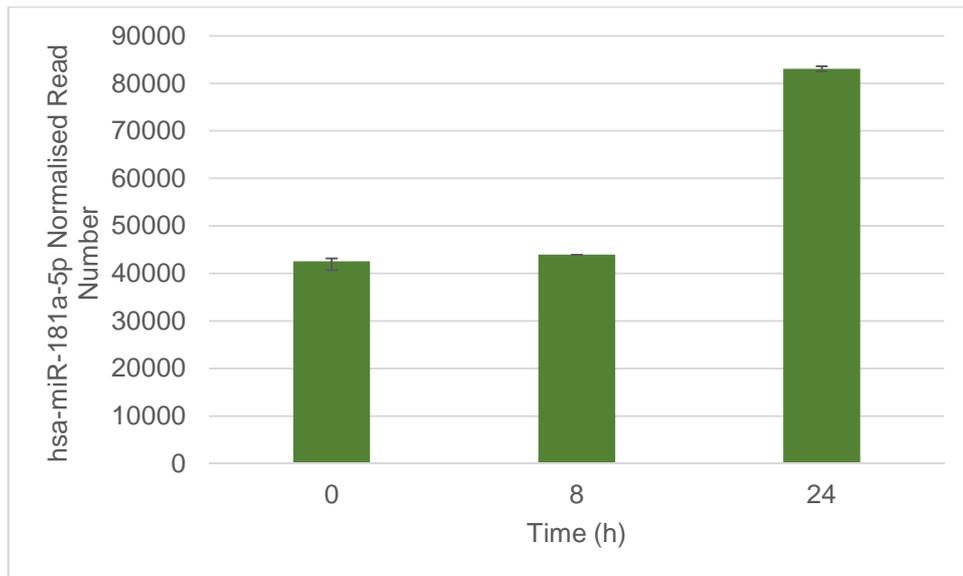


Figure 58: Normalised read numbers for hsa-miR-181a-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

It was initially speculated that miR-181a-5p modulation might have been masked by cross-hybridisation of the anti-miR-181a-5p Northern Blot probe with the closely-related hsa-miR-181b-5p, but subsequent dot-blot experiments did not imply the potential for such cross-hybridisation to occur to a significant extent (see Appendix for data).

Finally, the library data analysis implied that miR-345-5p was upregulated by SFN (10 μ M) in CCD-841 2.5-fold at 24 h (as illustrated in Figure 59).

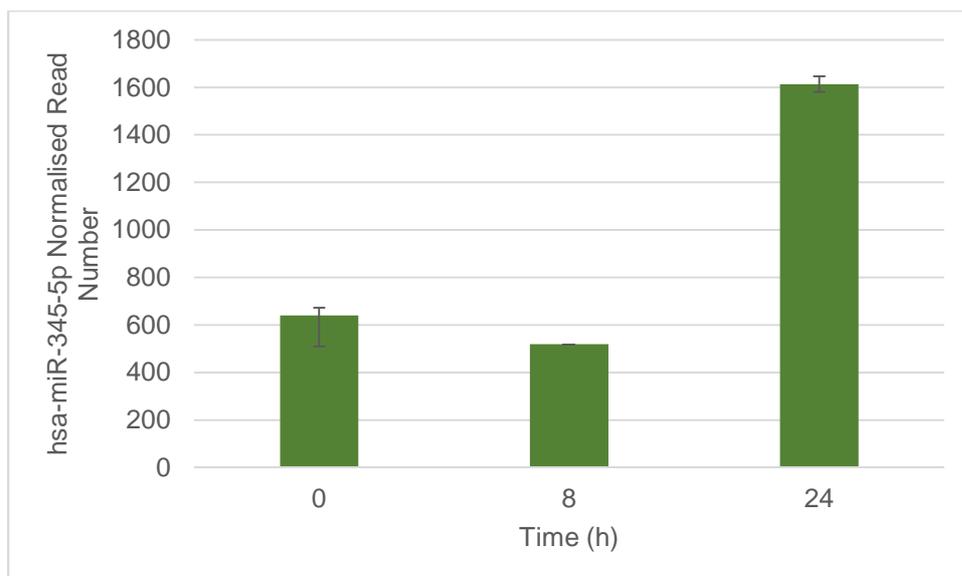


Figure 59: Normalised read numbers for hsa-miR-345-5p in CCD-841 cells following SFN treatment for 8 or 24 h, or in controls. Data are represented as medians across library replicates, with the maximum and minimum values of replicate sets represented as the upper and lower bounds of error bars.

However, subsequent RT-qPCR experiments did not corroborate this (see Appendix for data). In fact, the RT-qPCR data actually indicated downregulation by 20 μ M SFN at 24 h, which may have been due to mild cytotoxicity.

4.4 Summary

Caco-2 and CCD-841 cells were treated with SFN (10 μ M) for 8 or 24 h in triplicate, and then total RNA samples were extracted and used to clone miRNAs in libraries using the method and HD adapters developed by the Dalmay laboratory. The libraries were deep sequenced, and the sequencing data were quality-checked and normalised. Reads were then annotated against the list of known human miRNAs. Medians read numbers for each miRNA were taken across replicate libraries, and compared to identify apparently SFN-modulated miRNAs. Out of such miRNAs, those with apparent relevance to cancer and/or for which library reads were relatively high were selected for further study.

A number of cancer-related miRNAs were confirmed by either Northern Blots or RT-qPCR to be SFN-modulated in Caco-2 cells:

- Let-7f-5p was indicated by the library data as upregulated in Caco-2 by 8 and 24 h SFN treatment 1.4- and 2.1-fold respectively, and was shown by Northern Blot experiments to be upregulated 1.2- and 1.4-fold at 8 and 24 h respectively.
- Let-7g-5p was indicated by the library data as upregulated in Caco-2 by 1.3- and 2.0-fold at 8 and 24 h respectively, and was shown by the Northern Blot experiments to be upregulated 1.6-fold at 24 h.
- By contrast, neither let-7f-5p nor let-7g-5p were modulated by SFN treatment in CCD-841 according to the Northern Blot data.
- Dot-blot data suggested that the apparent increases of let-7f-5p and let-7g-5p in Caco-2 were not artefacts of probe cross-hybridisation with other let-7 family members.
- MiR-10a-5p[[isomiR]]21_42| was indicated by the library data as being upregulated 1.2- and 1.5-fold in Caco-2 by 8 and 24 h SFN treatment respectively, and also shown by the Northern Blot data to be upregulated 1.4-fold at 8 h, although not at 24 h.
- A Northern Blot experiment for miR-29b-3p showed that it was downregulated 1.4 and 1.6-fold in Caco-2, by 8 and 24 h SFN treatment respectively.
- MiR-193b-3p was indicated by the library data as being downregulated in Caco-2 by 8 and 24 h SFN treatment 1.5- and 2.0-fold respectively, and shown to be downregulated 1.4-fold at 8 h but not at 24 h, by the Northern Blot data.

Chapter 5: MicroRNA Target Validation

5.1 General Introduction

The two-step miRNA profiling process, i.e. library construction-sequencing-analysis followed by individual assays, has confirmed that SFN treatment modulates the expression of several reportedly cancer-related miRNAs in Caco-2 cells. Since miRNA-mRNA base pairing need not be perfect for the repression of translation, any given miRNA potentially regulates many different mRNA transcripts. Computational algorithms such as miRanda are often used to predict which mRNAs a given miRNA will bind to, based on factors such as miRNA seed region-mRNA 3'-UTR complementarity and binding energy. However, such algorithms inevitably make many incorrect predictions when loose parameters are used, and/or fail to predict interactions that actually occur if stricter parameters are input. Actual miRNA-mRNA interactions cannot thus be conclusively demonstrated by computational predictions alone; a cell-based assay such as the luciferase assay is necessary to confirm the ability of a given miRNA to bind (a) predicted mRNA target site(s).

5.2 Let-7f-5p

Analysis of the library data implied that let-7f-5p was upregulated by SFN (10 μ M) treatment in Caco-2 cells, and was highly abundant. This modulation was then confirmed by Northern Blots (as shown in Chapter 4).

Let-7 miRNAs, such as let-7f-5p, have been repeatedly demonstrated to have anti-proliferative, pro-differentiative and pro-apoptotic effects, thereby acting as tumour suppressors. Interestingly, let-7 miRNAs have been demonstrated to repress the translation of Bach1: a protein that competes with Maf proteins for binding to Nrf2 in the nucleus and thus represses Nrf2-mediated gene activation. Nrf2-Maf heterodimers are more transcriptionally activating of ARE-controlled genes than are Bach1-Maf heterodimers. In fact, let-7 miRNAs reportedly upregulated the Nrf2-activated reducing enzyme HO-1 via Bach1 repression (Hou, Tian et al. 2012). The potential for a given

miRNA to repress tumour cell growth in spite of the cytoprotective Nrf2-mediated antioxidant response is very interesting from a therapeutic point of view.

Let-7f-5p was initially selected for further study, based on its apparent high abundance and high likelihood of influencing carcinogenesis. Of interest are its effects at the mRNA and protein levels, as well as consequent effects on cancer 'endpoints' *in vitro* such as cell migration and invasion.

5.2.1 Selection of Predicted Target Genes

As discussed previously, algorithms that predict the mRNA binding sites of a miRNA often return thousands of predictions, many of which are likely to be false positives. 5467 genes were predicted as targets of let-7f-5p by the miRanda algorithm (<http://www.microrna.org>). Amongst those with mirSVR scores ≤ -1.2 , several were selected for validation and further study after scouting the literature to ascertain their apparent relevance to cancer (particularly CRC), and existing reports of interactions between them and let-7 miRNAs. Thus selected were:

- BACH1 – a repressor of Nrf2-mediated gene transcription, but also reported to promote CRC proliferation (Yun, Jo et al. 2013).
- CDC25A – a phosphatase that promotes cell cycle progression, and is reported to be overexpressed in many cancers (Boutros, Lobjois et al. 2007).
- HMGA2 – a promoter of EMT, proliferation, motility and invasion in CRC (Li, Zhao et al. 2014).
- KRAS – a growth promoter that is mutationally over-activated and/or overexpressed in many cancers, including CRC.
- MYC - a promoter of proliferation, migration and invasion that is mutationally over-activated and/or overexpressed in many cancers, including CRC.

5.2.2 Luciferase Assays

The 3'-UTRs of the genes of interest, as listed above, were cloned into the pmiRGLO vector as described in Chapter 2, for use in luciferase assays to

detect the binding of miRNAs to the 3'-UTRs. Variations of the 3'-UTRs produced by mutating the predicted miRNA-binding sites were also cloned into the vector as described in Chapter 2, to use as controls (see Appendix for plasmid verification and sequencing data). 'Positive control' experiments were performed to confirm that the let-7f-5p mimic and inhibitor were actually decreasing and increasing let-7f-5p functionality, respectively. For these experiments, an artificial insert containing four repeats of sequence antisense to that of let-7f-5p was created, and inserted into pmiRGLO vector in a similar fashion to the 3'-UTR clones.

HCT116 cells were co-transfected with insert-containing pmiRGLO vector (200ng) and let-7f-5p mimic (100nM) (Qiagen) and/or let-7f-5p inhibitor (100nM) (Qiagen) and/or AllStars Negative Control siRNA (100-200nM) (Qiagen), equalising 'RNA' concentrations to 200nM across all transfected wells. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit (Promega). Relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring luminescence. [Firefly luminescence]/[*Renilla* luminescence] ratios were calculated for each well, and then normalised to controls (no mimic or inhibitor; AllStars Negative Control siRNA at 200nM).

5.2.2.1 Control Experiments

As illustrated by the data in Figure 60, the control experiments confirmed the expected alteration of let-7f-5p functionality by the let-7f-5p mimic and inhibitor. The data show both a dramatic repression of relative Firefly luciferase expression by the mimic alone (69%), and a strong upregulation of such by the inhibitor alone (195%). As expected, the effects of the mimic neutralised those of the inhibitor, and vice versa.

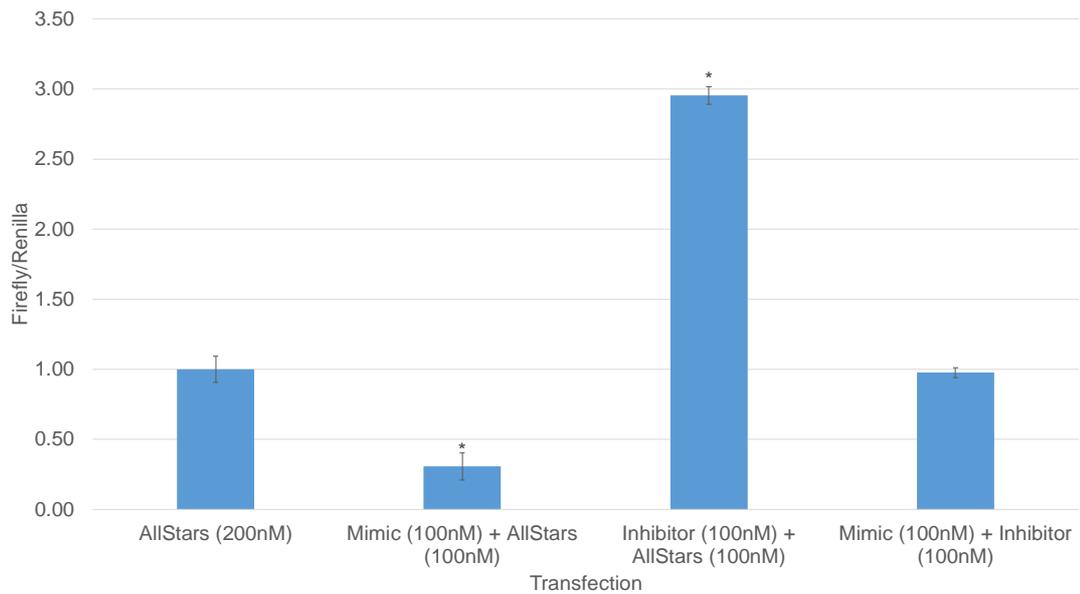


Figure 60: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insert consisting of four repeats of sequence antisense to that of let-7f-5p. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

5.2.2.2 BACH1

As illustrated by the data presented in Figure 61, the let-7f-5p mimic appears to inhibit relative Firefly luciferase expression from the vector containing the 'wild-type' form of part of the BACH1 3'-UTR, by 27% vs. control (albeit below significance; $p = 0.07$, according to unpaired two-tailed Student's T-test), but has no apparent effect on expression from the vector containing the 'mutated' form of the BACH1 3'-UTR segment. This may indicate that let-7f-5p actually binds to one or both of the predicted binding sites in its 3'-UTR, but a repetition of the experiment would be necessary to confirm this, due to failing to reach statistical significance. The let-7f-5p inhibitor had no obvious effect alone, but in combination with the mimic, abrogated the mimic's effect.

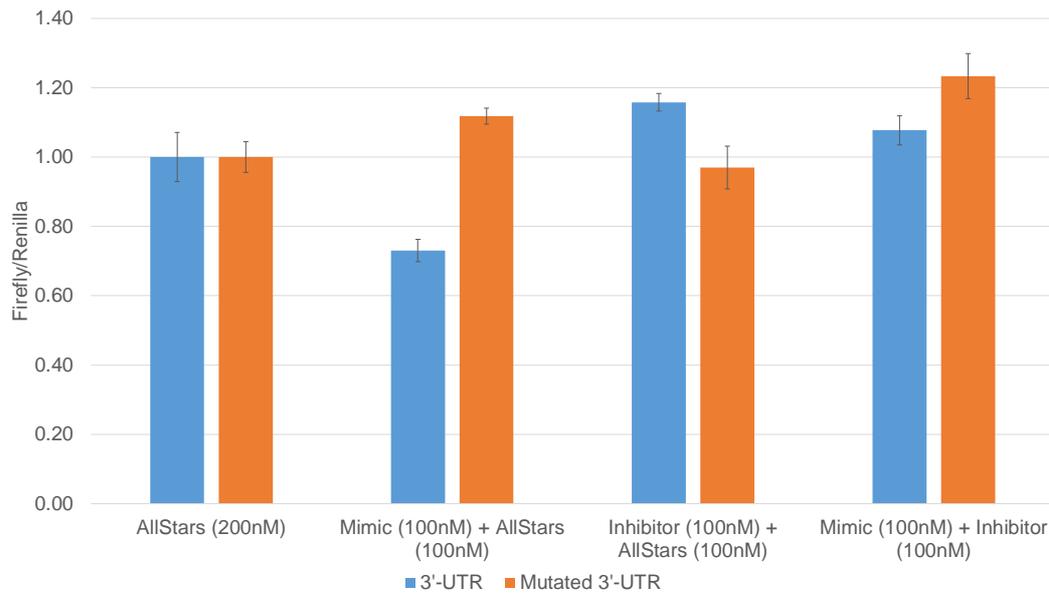


Figure 61: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the BACH1 3'-UTR containing two predicted let-7f-5p binding sites, or a modified form of this 3'-UTR portion with the predicted binding sites mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

5.2.2.3 CDC25A(a)

The data presented in Figure 62 indicate that although the mimic alone had no effect on relative Firefly luciferase expression, the inhibitor alone significantly increased such expression from the 'wild-type' vector by 43% ($p=0.04$), but not from the mutant vector. This implies that let-7f-5p is able to bind to the predicted site in the 3'-UTR of CDC25A. The combination of the mimic with the inhibitor did not abolish the inhibitor's upregulatory effects. The lack of observable effect of the mimic alone potentially implies that at the basal level of let-7f-5p in HCT116, the binding site is already saturated with respect to bound let-7f-5p.

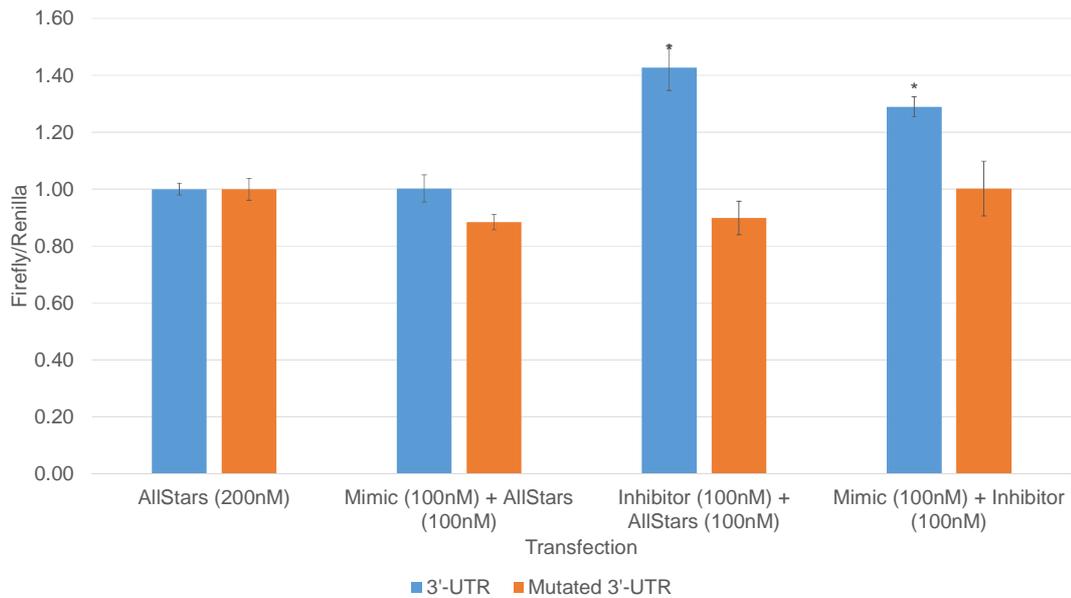


Figure 62: Data from luciferase assays performed after co-transfecting HCT116 cells with pmIRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmIRGLO vector contained an insertion of either a part of the CDC25A 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

5.2.2.4 HMGA2(b)

It is apparent from the data presented in Figure 63 that the let-7f-5p mimic significantly repressed relative Firefly expression from the 'wild-type' HMGA2 vector containing three predicted let-7f-5p binding sites, by 36% ($P=0.02$), but had no effect on that from the vector with mutated binding sites. The inhibitor alone appeared to slightly de-repress expression, but not to statistical significance ($P=0.13$), although the combination of the inhibitor with the mimic significantly abrogated mimic-induced repression. This confirms that one or more of the three predicted binding sites are bound by let-7f-5p in a transcription-repressing fashion. The lack of observable effect from the

inhibitor alone potentially implies that the binding sites are mostly unoccupied by let-7f-5p anyway at basal levels.

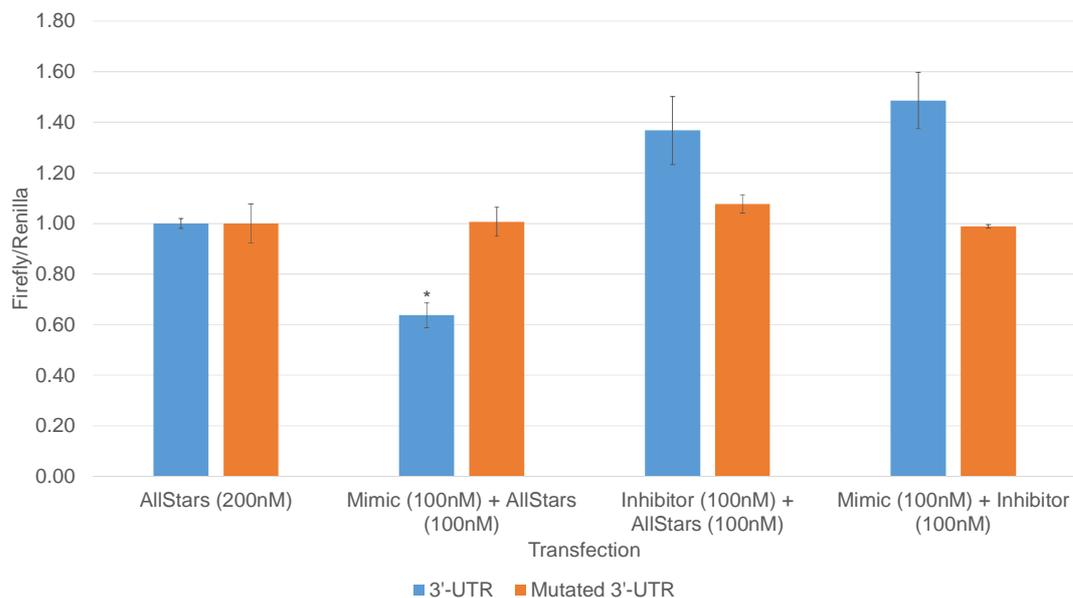


Figure 63: Data from luciferase assays performed after co-transfecting HCT116 cells with pmiRGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmiRGLO vector contained an insertion of either a part of the HMG2A 3'-UTR containing three predicted let-7f-5p binding sites, or a modified form of this 3'-UTR portion with the predicted binding sites mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

5.2.2.5 MYC

The let-7f-5p mimic alone did not impact relative Firefly luciferase expression from the MYC 'wild-type' vector, according to the data presented in Figure 64. However, the inhibitor alone increased expression from the 'wild-type' vector by 35% ($P=0.03$) but not that from the mutant vector. The transfection of the mimic together with the inhibitor reduced the inhibitor's effect. Surprisingly, the co-transfection of mimic and inhibitor dramatically increased relative Firefly luciferase expression from the mutant vector, although not from the 'wild-type' vector. Neither the mimic nor inhibitor had any such effect

alone. Perhaps the new sequence created by the artificial mutation was transcriptionally upregulated by this particular combination of mimic and inhibitor via unknown mechanisms.

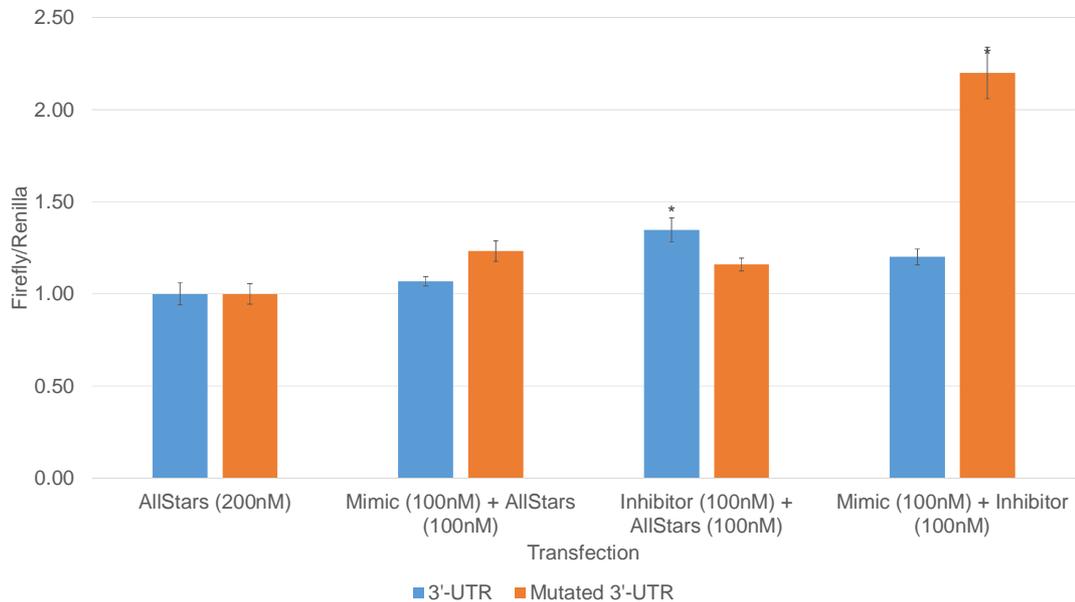


Figure 64: Data from luciferase assays performed after co-transfecting HCT116 cells with pmirGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmirGLO vector contained an insertion of either a part of the MYC 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

No effects of the let-7f-5p mimic or inhibitor on relative Firefly luciferase expression from the 'wild-type' KRAS vector are observable from the data illustrated in Figure 65. However, the mimic unexpectedly upregulated expression from the mutant vector only. The mutation had been generated by the creation of a 'GATATC' sequence (EcoRV restriction site). The 'GATATC' sequence *per se* is unlikely to have been solely responsible for the mimic-induced upregulation, because such upregulation was not observed with the mutant BACH1, CDC25A(a) and HMGA2(b) vectors, for

which the 'GATATC' motif was also introduced. Perhaps the mutation in this specific context of surrounding bases facilitated indirect transcriptional upregulation by let-7f-5p mimic via unknown mechanisms. Interestingly, the combination of let-7f-5p inhibitor with the mimic significantly diminished upregulation vs. the mimic alone ($p=0.001$).

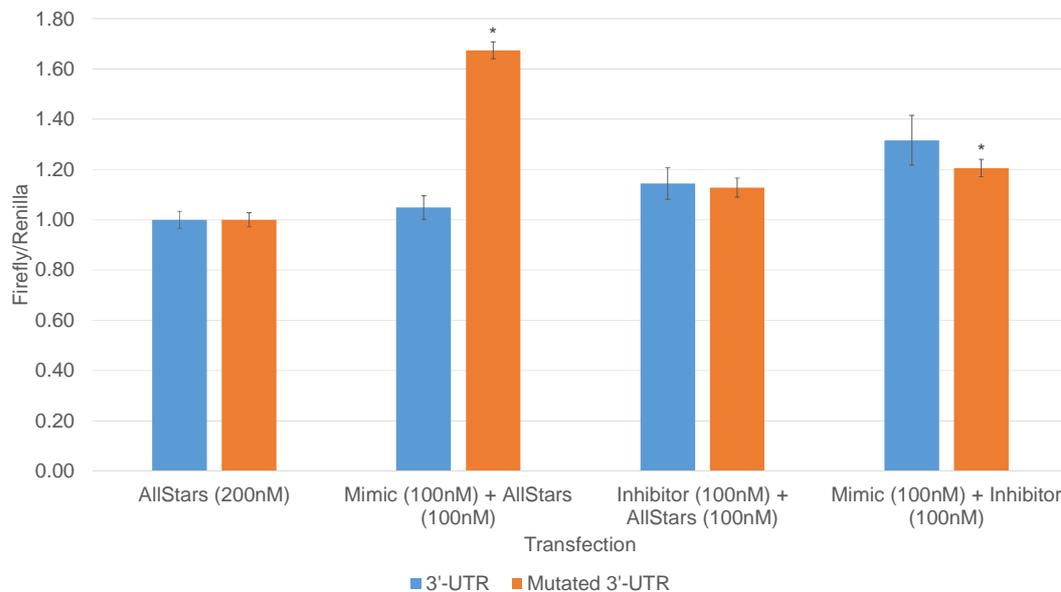


Figure 65: Data from luciferase assays performed after co-transfecting HCT116 cells with pmirGLO vector and let-7f-5p mimic (Qiagen) and/or let-7f-5p inhibitor (Qiagen) and/or AllStars Negative Control siRNA (Qiagen). The pmirGLO vector contained an insertion of either a part of the KRAS 3'-UTR containing a predicted let-7f-5p binding site, or a modified form of this 3'-UTR portion with the predicted binding site mutated. Luciferase assays were performed at 48 h post-transfection using the Dual-Glo Luciferase Assay Kit, and then relative Firefly and *Renilla* luciferase expression levels were evaluated by measuring the luminescence of metabolised substrates. Blank-corrected Firefly luminescence values were divided by blank-corrected *Renilla* luminescence values, and then ratios were normalised to controls (AllStars Negative Control siRNA at 200nM). Normalised data are represented as triplicate means \pm S.E.M. (* $p < 0.05$ according to two-tailed Student's T-test).

5.3 Conclusions

It is apparent from the data presented in this Chapter that let-7f-5p is able to bind to sites in the 3'-UTRs of CDC25A, HMGA2 and MYC predicted as binding sites. Therefore, it would be informative to perform experiments in the future to:

- Assay the effects of let-7f-5p mimic and/or inhibitor on the expression of these genes at the mRNA and protein levels.
- Determine whether or not sub-cytotoxic doses of SFN directly modulate such expression.
- Test for synergistic or antagonistic interactions between SFN and let-7f-5p mimic or inhibitor.

Chapter 6: General Discussion

6.1 Recapitulation of Study Rationale and Objectives

The risk of CRC, an increasingly important cause of morbidity and mortality globally whose symptoms tend not to become apparent until the disease has advanced to a terminal stage, is apparently linked to diet and lifestyle.

Interest in the anti-cancer potential of phytochemicals from fruits, vegetables and herbs, has intensified in recent decades. Given the inconclusiveness of studies into the relationship between overall fruit and vegetable intake and cancer incidence, yet observation of marked inverse correlations between the consumption of cruciferous vegetables and cancer incidence, interest in the potentially protective roles of ITCs from these vegetables has gained traction in recent decades. CRC is a particularly interesting target, since the colorectal tissues are inevitably better predisposed than many others to be exposed to ingested ITCs prior to their metabolism. Chemopreventative properties of ITCs have classically been attributed to hormetic Nrf2-mediated antioxidant effects, by which low-to-moderate doses tend to be cytoprotective, whilst high doses are potentially cytotoxic, although several studies indicate that the plasma concentrations reasonably attainable by cruciferous vegetable consumption are well below the cytotoxic threshold for healthy cells.

The roles of Nrf2 in cancer are very complex, since its cytoprotective effects in healthy cells are usually beneficial, whereas its activity in tumour cells is often considered undesirable because it may increase their chemoresistance. On the other hand, tumour cell Nrf2 activation can in some cases have counter-intuitive anti-proliferative and anti-metastatic effects. The interactions between ITCs and carcinogenesis clearly extend beyond Nrf2 activation, such as to the direct inhibition of HDACs.

MiRNAs are small non-coding RNAs that post-transcriptionally modulate gene expression, whose importance in human health (particularly in cancer) has gained increasing recognition over the past 15 years. It is also increasingly apparent that dietary, lifestyle and environmental factors can

influence miRNA expression levels, which themselves could be important mediators of the effects of such factors on health. The existence of evidence that SFN may modulate miRNA expression, the known ability of miRNAs to modulate carcinogenesis, and the attribution of ITCs with chemopreventative and/or chemotherapeutic effects, together form the rationale for investigating the potential roles of miRNAs in the anti-cancer effects of SFN, with a view to the development of more effective and/or less toxic chemotherapeutic strategies. CRC adenocarcinoma Caco-2 and non-cancerous colonic epithelial CCD-841 cells were used for the profiling of miRNA expression following SFN treatment (after an initial experimental validation of the system), which included a wide-scale miRNA expression analysis via deep sequencing, followed by individually-targeted miRNA assays for verification.

6.2 General Summary of Findings

6.2.1 Validation of the System

6.2.1.1 Cell Viability Assays

The cytotoxicity of SFN towards the Caco-2 and CCD-841 cell lines was evaluated, by using the MTT assay to measure cell viability at 24 or 48 h following the addition of SFN to the medium of cells in the logarithmic growth phase. Cytotoxicity was clearly demonstrated towards both lines at high doses, with IC₅₀ values of ~ 40µM at both 24 and 48 h in Caco-2, and IC₅₀[24h] and IC₅₀[48h] of 97 and 20µM in CCD-841. It was concluded that concentrations ≤10µM were not cytotoxic and thus were suitable for experimental use.

The cytotoxicity of a source of ROS, H₂O₂, was tested against both cell lines, giving IC₅₀[24h] and IC₅₀[48h] values of 200 and 320µM in Caco-2, and IC₅₀[24h] and IC₅₀[48h] values of 200 and 70µM in CCD-841. The data clearly showed that Caco-2 cells were able to recover from H₂O₂-mediated inhibition between 24 and 48 h, whilst CCD-841 cells were not, suggesting important differences between the cells that may be due to differential apoptotic and/or cell cycle arrest signalling.

6.2.1.2 The Effects of Sulforaphane on Nrf2 and Nrf2-Mediated Gene Expression

A clear SFN-induced upregulation of TrxR-1 at both the mRNA and protein level was demonstrated in both cell lines, according to RT-qPCR and Western Blot assays. Induction at the mRNA level was dose-dependent at 24 h (SFN \leq 20 μ M) in both cell lines, but the effects on the actual levels of this selenoprotein (at 48 h) were highly dependent upon the availability of selenium. Protein-level upregulation in the presence of supplemented selenium (in the form of SeMSC (0.2 μ M)) was apparently maintained for at least 72 h. Another Nrf2-activated reducing enzyme, UGT1A, was clearly demonstrated to be upregulated by SFN in a dose-dependent fashion, at both 24 and 72 h, in Caco-2.

A dramatic increase in nuclear Nrf2 levels was demonstrably induced by SFN in Caco-2 cells from 4 h. Such increases in nuclear Nrf2 did not appear to be dose dependent (5-20 μ M SFN), implying that the dose-dependent effects on TrxR-1 mRNA transcript levels may point to the influence of pathways complementary to Nrf2 nuclear translocation, such as the activation of Nrf2 co-factors and/or the inhibition of competitive inhibitors of Nrf2-ARE binding e.g. Bach1 (Reichard, Motz et al. 2007).

In conclusion, the data confirmed the ability of sub-cytotoxic doses of SFN to induce the nuclear translocation and transcription-activating activity of Nrf2 in the experimental set-up used, implying that such conditions would be appropriate for miRNA-profiling experiments.

6.2.2 Effects of Sulforaphane on Oncogenic Characteristics

6.2.2.1 Cell Migration: Wound Healing Assays

The wound healing assay assesses the migration of cells into a linear gap 'wound' created by scratching a confluent cell monolayer, by measuring the average distance between leading cells down the gap axis after a fixed period of time post-scratch (usually 48 h). In this instance, treatment-induced repression of migration was inferred from increases in endpoint average gap width in treated (SFN) vs. control (DMSO) monolayers. SFN treatment clearly inhibited CCD-841 cell migration in a dose dependent fashion (SFN \leq

20 μ M), at concentrations demonstrated to be sub-cytotoxic by MTT assays. It is unclear whether or not this implies that SFN can impair the healing and renewal of non-cancerous colonic tissue. It might be expected that early tumour cells would probably react in a similar manner. No effects on migration were apparent in the Caco-2 cell line, which is derived from an advanced colorectal adenocarcinoma. It could be postulated that Caco-2 cells are resistant to SFN-mediated repression of migration due to mutations and/or epimutations. On the other hand, the highly 'adhesive' nature of Caco-2 cells may have prevented them from migrating whether or not SFN was present. A future experiment in a different CRC cell line, such as HCT116, might be more informative.

6.2.2.2 Clonogenic Assays

The clonogenic assay measures the ability of cells to divide and form colonies, following seeding at densities so low that cells are isolated from one another. The ability to survive and divide many times in spite of isolation is characteristic of highly oncogenic cells.

Cells were seeded at very low density, and then SFN was added to culture medium at different concentrations. Cultures were then incubated for 2 weeks, at the end of which the numbers of colonies (clusters of ≥ 50 cells) formed in SFN-treated wells were compared to those in control (DMSO only added) wells. This showed a very clear and dose-dependent inhibition of Caco-2 clonogenicity in response to the addition of SFN to culture medium, by 50 and 90% at 5 and 20 μ M SFN respectively. However, it was considered that these effects were not necessarily independent of SFN-induced oxidative stress and/or genotoxicity, because the addition of SFN to the culture medium of cells present at very low density results in much greater per-cell SFN exposure than if cells are present at 70% confluence.

Therefore, another experiment was done in which cells were grown to 70% confluence, then pre-treated with the same concentrations of SFN for 24 h, and then detached and replated at low density for 2 week incubation without any further addition of SFN. Such acute sub-cytotoxic dose pre-treatment had no discernible effect on colony formation.

6.2.3 Profiling MicroRNA Expression Following Sulforaphane Treatment

Conditions similar to those in which SFN-mediated Nrf2 and antioxidant gene induction was demonstrated were used to treat Caco-2 and CCD-841 cells, and then total RNA was extracted for miRNA profiling. Cells were treated for 8 or 24 h with SFN at a concentration of 10 μ M.

6.2.3.1 MicroRNA Libraries

MiRNA libraries were constructed from total RNA samples using HD adapters (Sorefan, Pais et al. 2012) and the cloning method (Xu, Billmeier et al. 2015), both developed by the Dalmay laboratory. The main purpose of miRNA library construction, sequencing and analysis was to highlight apparently SFN-modulated cancer-related miRNAs upon which further studies should be focussed. Constructed libraries were deep sequenced on the Illumina HiSeq 2500 platform by the Earlham Institute, Norwich (formerly The Genome Analysis Centre). Deep sequencing data were quality-checked, reads were annotated against the list of known human miRNAs using PatMan, and then annotated data were manually scrutinised to check for misannotations. Apparent fold-changes in miRNA expression were determined by comparing normalised read numbers for each miRNA between the different libraries.

Analysis of the data revealed the presence of isomiRs, varying in sequence with respect to the canonical forms by having several missing or extra nucleotides at either end. In most cases, the canonical form of a given miRNA dominated the read population, but alternative isomiRs dominated in a number of cases. IsomiR dominance may be influenced by cell and tissue type, and may also vary between tumour cells and their non-cancerous counterparts in a given tissue (Tan, Chan et al. 2014).

6.2.3.2 MicroRNAs that were Individually Confirmed to be Modulated

Of miRNAs indicated to be modulated by SFN according to the library sequencing data, several were selected for validation, including: let-7f-5p, let-7g-5p, miR-10a-5p, miR-10b-5p, miR-193b-3p, miR-182-5p, miR-17-3p,

miR-1296-5p, miR-106b-5p, miR-181a-5p and miR-345-5p. Additionally, miR-29b-3p was individually assayed.

The let-7 miRNAs are frequently reported as tumour suppressors by promoting cell cycle arrest, differentiation and/or apoptosis, and inhibiting proliferation. Interestingly, several let-7 miRNAs have been demonstrated to repress Bach1, which is a competitive inhibitor of Nrf2-ARE binding (Hou, Tian et al. 2012). Let-7f-5p and let-7g-5p both appeared to be highly abundant and upregulated by SFN in Caco-2 according to the library data; these upregulations were subsequently corroborated by Northern Blot experiments. For the sake of comparison, Northern Blots were performed to test the effects of similar SFN treatment in CCD-841, in which interestingly, SFN appeared to have no effect on the expression of let-7f-5p or let-7g-5p. In order to rule out the possibility of the Northern Blot data for these miRNAs being confounded by non-specific cross-hybridisation of the probes, oligomers composed of sense let-7 miRNA sequences were blotted onto a membrane, and then the anti-let-7f-5p and anti-let-7g-5p Northern Blot probes were tested to see if they would cross-hybridise with the other let-7 family members. The dot-blots indicated very little-to-no potential for cross-hybridisation.

MiR-10a-5p was also indicated to be upregulated by SFN in Caco-2 according to the library data, and this was corroborated by Northern Blot and RT-qPCR experiments. The library data conversely indicated miR-193b-3p to be downregulated by SFN in Caco-2 at 8 and 24; a Northern Blot corroborated reduced expression at 8 h but not at 24 h. Given that the library data exhibited high inter-replicate variability, the Northern Blot data are likely to be more representative of reality.

Interestingly, miR-29b-3p was shown to be downregulated by SFN in Caco-2 according to a Northern Blot. This had not been highlighted by the library data analysis, most likely because the normalised read numbers for this miRNA had wide inter-replicate variation. The downregulation of miR-29b-3p could have been mediated by Nrf2-mediated repression of transcription from

the miR-29b-1 locus, which has been demonstrated in acute myeloid leukaemia cells (Shah, Zaitseva et al. 2015).

6.2.3.3 MicroRNAs Unable to be Validated and Possible

Explanations

However, the apparent modulations of some other miRNAs were not corroborated by Northern Blots and/or RT-qPCR experiments, particularly in CCD-841. Based on the data for miRNAs individually assayed in CCD-841 thus far, it is apparent that the CCD-841 library data are unlikely to reflect reality. Several significant fold-changes at 24 h were apparent from the CCD-841 library data, but were not corroborated by the Northern Blot and/or RT-qPCR experiments, such as for miRNAs including miR-181a-5p, miR-182-5p and miR-345-5p. A potential contributor to these discrepancies is the fact that the total RNA samples used for 24 h library construction were extracted in a separate instance (different experiment) from those samples used for the construction of the other libraries. Cells used in the second instance had been treated under the same conditions as those used in the first, but were at a different passage number. At the time, this was inevitable because the previously-extracted 24 h total RNA samples had been used up in previous library construction attempts.

For future miRNA library construction-sequencing-analysis experiments, it would be advisable, where possible, to ensure that all samples used for library construction are extracted in the same experiment. This should help to minimise inter-replicate variation with respect to normalised read numbers. Therefore, preferably large quantities of RNA ought to be extracted at once, so that unsuccessful attempts at library construction can be repeated using the same RNA samples. This was taken into account when carrying out subsequent experiments for the Caco-2 libraries, all of the RNA samples used for which were extracted in the same experiment.

At the time of CCD-841 library construction, it was not expected that RNA samples extracted from a separate experiment would exhibit wildly different miRNA expression profiles if similar experimental conditions were maintained. However, given that relatively high inter-replicate variation is

already observable across normalised read numbers for some miRNAs even in the Caco-2 library data, it is prudent to minimise unnecessary potential sources of variation. As to the question of why such inter-replicate variation was observed in the Caco-2 library data, it is possible that the long RNase-sensitive nature of the 5'- and 3'-adapter ligation process facilitated some stochastic RNase-mediated miRNA degradation, which could perhaps be ameliorated by increasing the concentrations of RNase inhibitors in the reactions.

6.2.4 Validating Predicted MicroRNA Targets and Functions

6.2.4.1 Predicted MicroRNA-Target Genes

Let-7f-5p was selected for initial further study, given the validation of its sequential upregulation by SFN treatment in Caco-2, its apparent high abundance, and the frequency with which it is reported as a tumour suppressor. The 'microRNA.org – Targets and Expression' database catalogues miRNA-mRNA 3'-UTR interactions predicted by miRanda. The predicted target mRNAs of let-7f-5p according to this database, with mirSVR score ≤ -1.2 , were searched in the literature for existing reports of relevance to cancer. On this basis, BACH1, CDC25A, HMGA2, KRAS and MYC were selected for validation.

There are limitations to such algorithms that currently make predictions based upon several factors including complementarity between the miRNA 5' seed region and mRNA 3'-UTR, binding energy and conservation across species. Inevitably, false positives arise since there are cases in which an miRNA and mRNA are not shown to actually bind in an experimental setting, despite strong or perfect 5' seed region-3'-UTR complementarity.

Conversely, algorithms can fail to predict actual miRNA-mRNA interactions, particularly those of a non-canonical nature that do not depend upon 5' seed region-3'-UTR complementarity. This has been evidenced by crosslinking, ligation and sequencing of hybrids (CLASH) experiments in HEK293 cells, which revealed the AGO-mediated targeting of mRNAs via miRNAs in the absence of miRNA 5' seed region-mRNA 3'-UTR complementarity (Helwak, Kudla et al. 2013).

6.2.4.2 Luciferase Assays to Confirm MicroRNA-3'-UTR

Interactions

The luciferase reporter assay was used to determine whether or not let-7f-5p was able to actually bind to computationally-predicted sites in the 3'-UTRs of the BACH1, CDC25A, HMGA2, KRAS and MYC transcripts. Portions of 3'-UTRs containing predicted let-7f-5p binding sites were cloned from human genomic DNA and then inserted into the pmiRGLO Dual Reporter vector (Promega), upstream of the Firefly luciferase coding region. Prepared vectors were co-transfected with let-7f-5p mimic and/or inhibitor into HCT116 cells. 'Mutant' forms of each vector were created in which the predicted let-7f-5p binding sites were destroyed. Additionally, a vector was prepared with an artificial insert based upon four repeats of sequence antisense to that of let-7f-5p, for initial 'control' experiments to confirm that the mimic and inhibitor were in fact impacting let-7f-5p function as intended. Assays using the 'mutant' and 'wild-type' vectors were run in parallel, in order to determine whether or not the predicted binding sites were actually responsible for the observed mimic and/or inhibitor-induced transcriptional regulation.

The initial 'control' luciferase assay confirmed that the let-7f-5p mimic and inhibitor did indeed increase and decrease let-7f-5p functionality, respectively. Data from subsequent luciferase assays confirmed that let-7f-5p was able to interact with predicted binding sites in the 3'-UTRs of CDC25A, HMGA2 and MYC, and probably BACH1 (although the mimic-induced repression of relative Firefly expression from BACH1 did not reach statistical significance; $p=0.07$), to induce transcriptional repression.

Originally it was expected that if let-7f-5p was able to bind to a particular site, the mimic alone would repress relative Firefly luciferase expression from the 'wild-type' vector and the inhibitor alone would upregulate such expression. However, while the mimic alone significantly repressed Firefly luciferase expression from the HMGA2 'wild-type' vector, the inhibitor alone did not significantly upregulate it. Conversely, the mimic alone had no significant effect on expression from the CDC25A 'wild-type' vector, but the inhibitor alone significantly upregulated it. This implies that the dynamics of 3'-UTR-

let-7f-5p binding, with respect to let-7f-5p concentration, differ according to 3'-UTR. Hypothetically, the observed data could be explained if:

- The basal let-7f-5p concentration in HCT116 cells is such that the CDC25A sites are already saturated, while the HMGA2 binding sites are largely unoccupied.
- The addition of inhibitor reduces the effective let-7f-5p concentration below a critical threshold for CDC25A, above which the CDC25A let-7f-5p binding sites become saturated.
- The addition of mimic increases the effective let-7f-5p concentration above a critical threshold for HMGA2, above which the HMGA2 let-7f-5p binding site rapidly transitions from being largely unoccupied to being nearly or completely saturated with let-7f-5p.

This hypothesis is illustrated by the sketch in Figure 66.

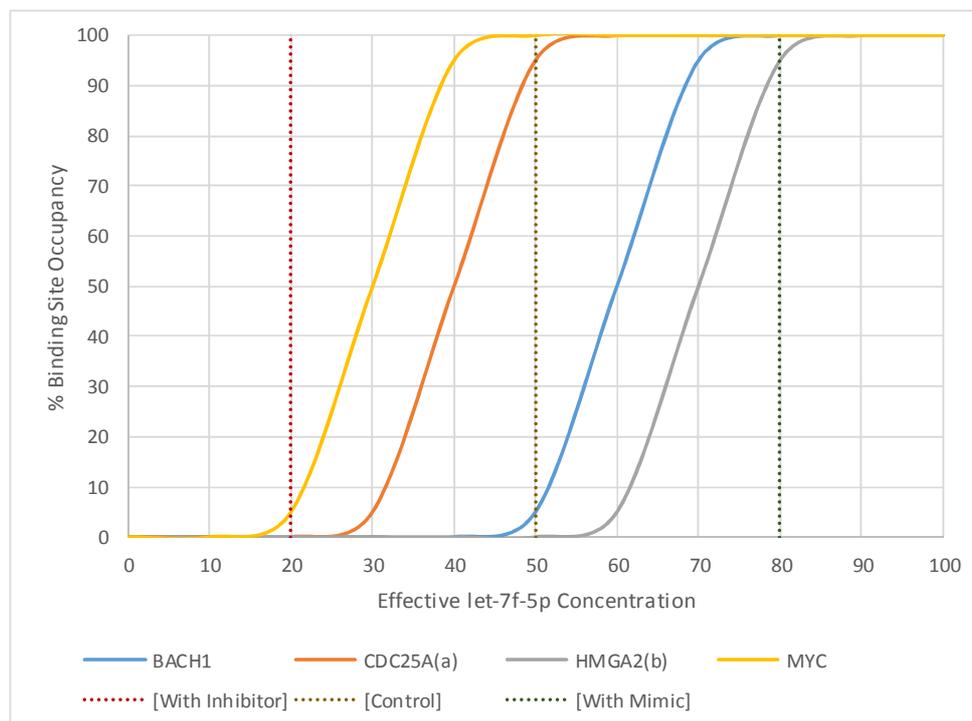


Figure 66: A sketch illustrating the hypothesis that the let-7f-5p site-binding dynamics are distinctly non-linear and differ between different 3'-UTRs. In the hypothetical situation here illustrated, the basal effective let-7f-5p concentration means that the CDC25A and MYC binding sites are already saturated with let-7f-5p but the BACH1 and HMGA2 sites are largely unbound. Thus, addition of let-7f-5p mimic has little to no effect on CDC25A and MYC, but strong effects on BACH1 and HMGA2, whilst the

addition of let-7f-5p inhibitor has little to no effect on BACH1 and HMGA2 but significantly decreases the occupancy of CDC25A and MYC sites by let-7f-5p.

No effects of let-7f-5p mimic nor of inhibitor on relative Firefly luciferase expression from the 'wild-type' KRAS vector were observed. However, an unexpected and strong upregulation of relative Firefly luciferase expression from the KRAS 'mutant' vector by the let-7f-5p mimic occurred, and this upregulation was significantly reduced by combining the inhibitor with the mimic ($p=0.001$). This suggests that the newly generated sequence by mutation was susceptible to let-7f-5p-induced upregulation by unknown mechanisms.

6.3 MicroRNA Library Construction: Technical Challenges

6.3.1 Initial Set of Caco-2 Libraries

An initial set of Caco-2 miRNA libraries was constructed using the ScriptMiner™ Small RNA-Seq Library Preparation Kit (Epicentre), rather than the individual reagents and improved protocol later developed by Xu et al. (2015). This initial set of libraries was unsuccessful in terms of there being several issues with the deep sequencing data. Firstly, very high proportions of reads in each library (50-85%) were 'wasted' because they were of 5'-adapter-3'-adapter dimers (AADs). Secondly, of non-AAD reads, very low proportions in each library could be mapped to the human genome (1-44%; 15% across libraries overall). It is to be expected for 20-25% of reads not to map to the human genome when considering post-transcriptional modifications, reverse transcription/PCR artefacts, and/or sequencing errors. However, the fact that 85% of reads overall could not be mapped to the human genome was a major concern, and possibly indicated contamination with non-human RNA during the library construction process. Upon removing the non-genome-mapping reads, remaining reads varied widely between libraries (from 50 794 to 2 481 053), making them unsuitable for profiling differential miRNA expression.

6.3.2 New Sets of Libraries for Caco-2 and CCD-841

An improved library construction protocol was developed in the laboratory by Xu et al. (2015), based upon the ScriptMiner™ Small RNA-Seq Library Preparation Kit, but using individual reagents and including adjustments to the kit manufacturer's protocol to improve cloning efficiency (Xu, Billmeier et al. 2015). A second set of Caco-2 libraries was constructed using this improved protocol, in order to maximise the chances of generating useful data. This new protocol was also used for the construction of the CCD-841 libraries.

PAGE size separation and analysis performed prior to sequencing clearly showed that the formation of 5'-adapter-3'-adapter dimers (AADs) was drastically reduced by the new method of library construction, and that miRNAs were more efficiently captured by the adapters. Following the reverse transcription and PCR amplification of captured small RNAs, the products were size-separated by PAGE, so that those of cloned miRNAs were separated from those of other sRNA species and from AADs, and formed visible bands between 120 and 140 bp. Following PAGE, gel pieces represented by bands appearing between 120 and 140 bp were cut out, and the DNA was eluted out of the gel fragments then purified.

Following PAGE separation, the stained bands can look discrete by eye, but the actual size-separation of the nucleic acids is never perfect. Inevitably, excised gel pieces will be contaminated with at least some clones of other sRNA species and AADs, so PAGE-extraction serves to minimise, rather than completely eliminate, such contamination. Evidently, this was insufficient for the first set of libraries, because > 50% of the sequencing reads were of AADs. This was probably a result of inefficient adapter-RNA ligation producing low 'adapter-miRNA-adapter':AAD ratios. Although the PAGE-extraction would have discarded the vast majority of the AAD molecules, the remaining minority present in the excised gel pieces might still have exceeded the captured RNA sequences in number.

For instance, one might imagine a hypothetical situation in which 1000 AADs are generated for every 1 captured miRNA. Gel extraction discards 99% of

the AADs but 1% remain present in the visible “microRNA” bands and are eluted. 1% of 1000 is 10, so there would still be 10 times as many AAD molecules as captured miRNA molecules in the final library preparation. Thus, the formation of AADs during the preparation of the first set of libraries may have been so great that the gel extraction performed could not sufficiently remove them.

The ‘adapter-miRNA-adapter’:AAD dimer ratio was significantly increased by the new protocol, as apparently from the bands visualised in the PCR-PAGE gels. Example PAGE gel images for two Caco-2 libraries – one constructed via the old protocol and the other by the improved protocol – are shown in Figure 67.

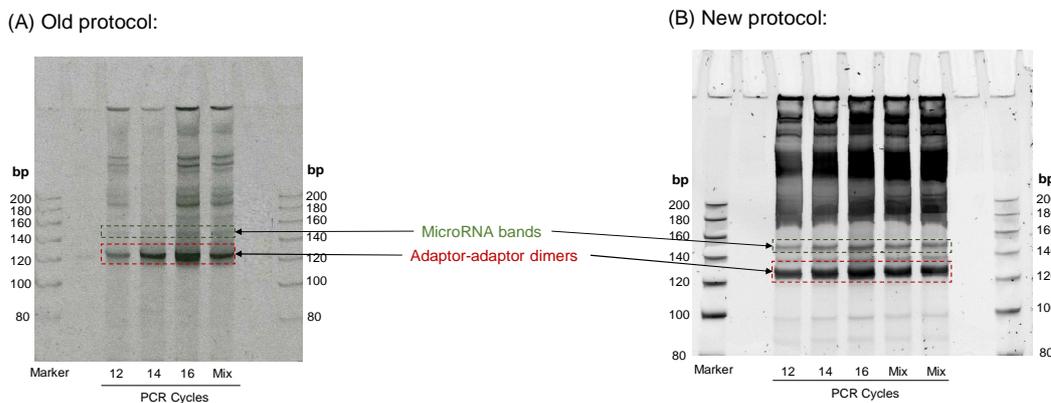


Figure 67: Images of SYBR Gold-stained PCR PAGE gels for Caco-2 libraries; one constructed by the old protocol (A), and the other by the new protocol (B). Although prominent AAD bands are apparent in both gels, their intensities relative to those of the higher miRNA bands are much lower in the gel image for the library constructed by the new protocol than in that for that constructed by the old protocol.

As expected, the sequencing data for the new libraries had far fewer AAD reads by proportion, which comprised only 0-30% and 0-36% of total reads from the new Caco-2 and CCD-841 libraries respectively.

6.4 MicroRNA Library Data vs. Findings in Literature

6.4.1 Inconclusive Observations

MiR-27b-5p was reported by Slaby et al. (2013) as upregulated by 48 h SFN treatment in NCM356 and NCM460 cells, but the CCD-841 and Caco-2 library data implied that SFN affected miR-27b-5p and miR-27b-3p expression neither at 8 nor 24 h. However, it is unknown whether or not modulation would have been observed in either cell line had expression been profiled at 48 h.

A number of miRNAs - miR-30a-3p, miR-505-3p and miR-629-5p - were found to be upregulated by 48 h SFN (10 μ M) treatment in the NCM356 and NCM460 non-cancerous colonic cell lines by Slaby et al. (2013), but indicated by the library data as being modulated by SFN (10 μ M) neither in CCD-841 nor in Caco-2 cells (time \leq 24 h). It cannot be ruled out that effects may have been observed had expression been profiled at later time points, however. Slaby et al. (2013) also found miR-758-3p to be upregulated in NCM356 and NCM460 cells by 48 h SFN treatment, but the CCD-841 library data highlighted no such modulation, whilst the Caco-2 library data had no reads for this miRNA.

6.4.2 Undetected MicroRNAs

A few miRNAs found to be differentially expressed by Slaby et al. (2013) did not have any reads at all in the Caco-2 and/or CCD-841 library deep sequencing data. These include miR-633, miR-106a-3p, miR-23b-3p, miR-145-5p and miR-372-5p/3p. The failure to detect these miRNAs could be due to several possible factors, including very low or no expression of these miRNAs in CCD-841 and Caco-2 cells. Another possibility is that there was a strong ligation bias against them during the miRNA cloning process. MiR-758-3p – upregulated by 48 h SFN treatment in NCM356 and NCM460 according to Slaby et al. (2013) – had reads in the CCD-841 library data, but was not indicated by such to be SFN-modulated. Reads for this miRNA were absent from the Caco-2 library data, however. Perhaps its expression has a tendency to be lost or highly downregulated in CRC, or is specifically absent in the Caco-2 cell line.

MiR-9-5p and miR-9-3p were reported by Slaby et al. (2013) to be upregulated in the non-cancerous NCM356 and NCM460 cell lines by 48 h SFN (10 μ M) treatment. However, the library data did not imply that either of these miRNAs were modulated by SFN in CCD-841 cells, although suggested slight upregulation in Caco-2 cells.

6.5 Future Perspectives

The data clearly show that sub-cytotoxic SFN treatment (at 10 μ M) upregulates the expression of let-7f-5p and let-7g-5p in the CRC cell line Caco-2, but not in the non-cancerous colonic cell line CCD-841. These data are not necessarily conclusive in demonstrating that SFN-mediated let-7 induction occurs exclusively in cancerous cells, but it would be interesting to test SFN-mediated let-7 induction in several other CRC and non-cancerous colonic cell lines, in order to observe any trends. Any such selectivity would be interesting and potentially very useful for limiting cytotoxicity towards non-cancerous cells; let-7 miRNAs are well-characterised for their pro-apoptotic and anti-proliferative effects in a raft of contexts, including in non-cancerous neural cells, in which such effects could promote neurodegeneration (Lehmann, Kruger et al. 2012).

The mechanisms by which SFN actually upregulates let-7 miRNA expression are unclear, but might involve SFN's HDAC-inhibiting activity, leading to the increased acetylation and thus transcriptional de-repression of the genetic loci that give rise to the let-7 miRNAs. It cannot be ruled out that yet-uncharacterised effects of SFN on miRNA-processing machinery and/or let-7 stability may have contributed to let-7 induction. The relationship between redox status and let-7 miRNA expression is unclear, although it has been reported that H₂O₂ can actually downregulate let-7g-5p expression in gastric cancer cells – a downregulation that promoted their resistance to X-ray-mediated cytotoxicity (Hu, Zhao et al. 2015). H₂O₂ is also reported to downregulate let-7a-5p and let-7b-5p in HCT116 cells, but surprisingly, in a p53-dependent fashion, whereby p53 represses transcription at loci giving rise to those miRNAs (Saleh, Savage et al. 2011). On the other hand, p53 is reported to positively regulate let-7 levels in PA1 human ovarian cancer cells

by inducing tristetraprolin, which downregulated the let-7 repressor, *lin28a* (Lee, Kim et al. 2013). Therefore, the links between SFN's acute pro-oxidant and/or Nrf2-inducing effects, and let-7 induction, are apparently complex and context-dependent.

As illustrated by the data presented in Chapter 3, It is clear that SFN significantly upregulates several cytoprotective phase II enzymes in both the tumour Caco-2 and non-tumour CCD-841 cell lines. Although such cytoprotective responses are generally considered to be undesirable in tumour (particularly advanced) cells, the net effects of such Nrf2-mediated responses on migration and/or invasion can counterintuitively be negative in some cases. As previously described, members of the let-7 family have been reported to upregulate the Nrf2-induced cytoprotective enzyme HO-1 in human hepatocytes via Bach1 repression, thus helping to protect against oxidative stress (Hou, Tian et al. 2012). Interestingly, Bach1 knockdown has been shown to upregulate TrxR-1 at the mRNA level in HepG2 cells (Hintze, Katoh et al. 2007), so it is conceivable that miRNA-mediated Bach1 repression is a contributing factor to SFN-mediated TrxR-1 induction.

It is possible that miRNAs such as those of the let-7 family may contribute to the apparent selectivity of SFN-mediated cytostatic and/or cytotoxic effects towards tumour cells vs. their non-cancerous counterparts that is frequently reported for given doses or concentrations used *in vitro*. Given that Nrf2-induced cytoprotective responses to SFN occur indiscriminately in both cancerous and non-cancerous cells, one might consider that a selective induction of let-7 miRNAs in the cancerous cells might help to counterbalance their Nrf2-induced cytoprotection, whilst sparing the non-cancerous cells from let-7-induced cytostasis and/or cytotoxicity.

6.5.1 Possible Future Experiments

6.5.1.1 Effects of let-7f-5p Mimic and Inhibitor on Target Expression

For those genes whose transcripts have been confirmed by luciferase assays to be bound in their 3'-UTRs by let-7f-5p, it would be a good idea to assay the effects of let-7f-5p mimic or inhibitor on protein-level expression by Western Blots, in order to confirm that the miRNA-3'-UTR binding actually downregulates translation. It would also be interesting to test for any effects at the mRNA level by RT-qPCR, although these will not necessarily occur for let-7f-5p-regulated genes. For target genes confirmed to be downregulated or upregulated at the protein level by let-7f-5p mimics or inhibitors respectively, the effects of SFN (5-20 μ M) could then be tested, alone and/or in combination with miRNA mimic or inhibitor, in order to test for antagonistic or synergistic effects. Additionally, the effects of specific protein-targeted siRNA knockdown on cancerous cell characteristics such as migration and invasion could be tested, in order to evaluate their significance with regards to SFN's and/or let-7f-5p's anti-tumour effects.

Data from these experiments could shed light on particular let-7f-5p-targeted oncogenes that are downregulated in CRC cells by sub-cytotoxic SFN exposure, and on the potential for such exposure to inhibit tumour progression despite Nrf2-induced cytoprotective responses in the tumour cells.

6.5.1.2 Other Potential Let-7 Targets and Sulforaphane-Modulated MicroRNAs

In the microrna.org database are listed miRanda-based predictions for a number of other mRNAs that are targets of let-7f-5p and/or let-7g-5p, with mirSVR scores ≤ -1.2 . These are described in Figure 68.

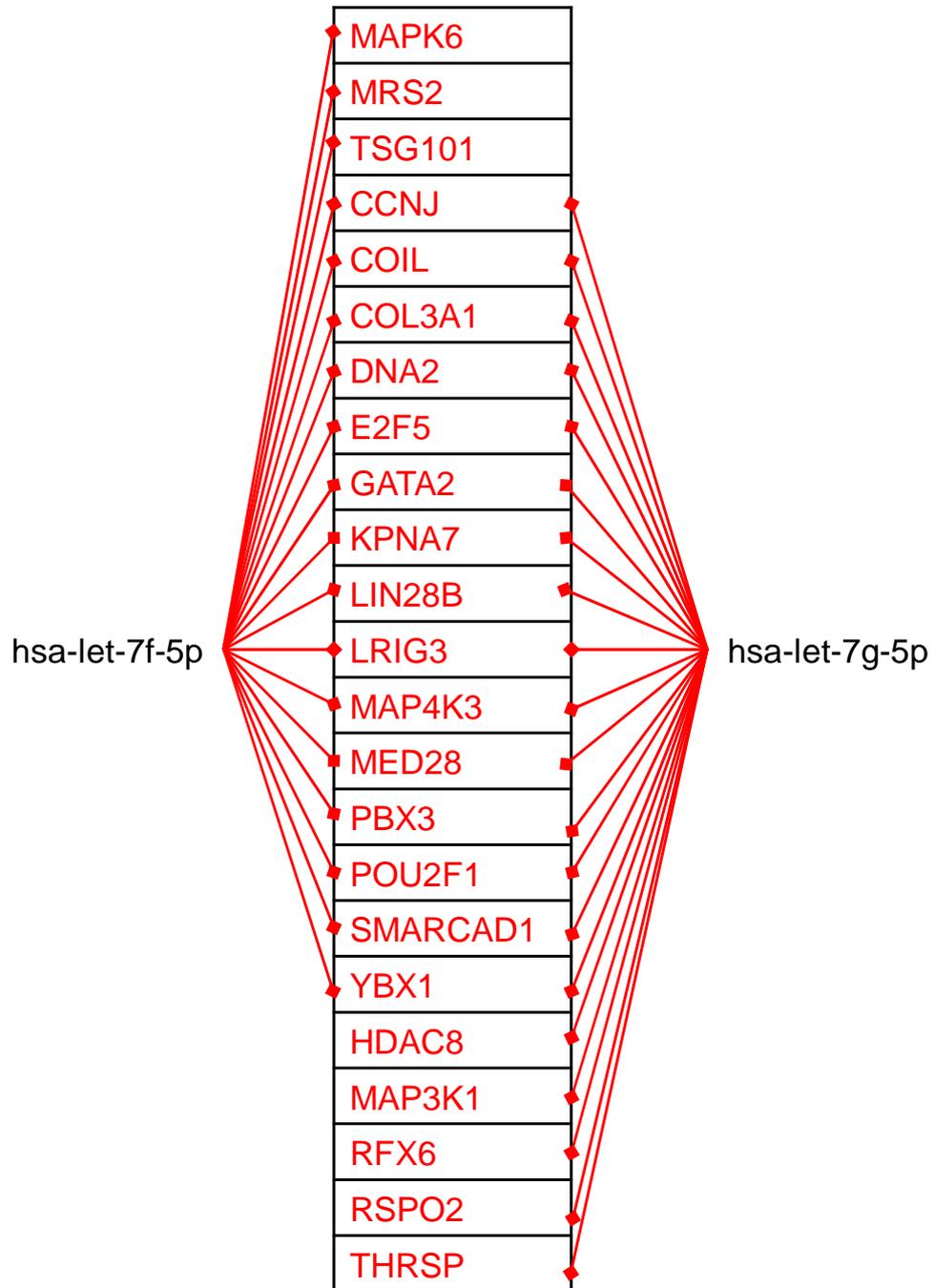


Figure 68: A diagram illustrating potentially oncogenic miRNAs that are predicted by miRanda (with mirSVR scores ≤ -1.2) to be targeted by let-7f-5p and/or let-7g-5p.

It would be interesting to validate some of these other target predictions by luciferase assays, and then further study their roles in SFN's anti-cancer effects. Considering that miR-10a-5p is also potentially upregulated by SFN, it may be relevant to also consider the potentially oncogenic miRanda-predicted target genes of that miRNA, including: ALCAM, BCL9, BDNF,

CNTN1, DYP24A1, DPAGT1, EIF4E, GALNT1, GATA6, PHF20L1, PIK3CA, SMC4 and WDR26.

Regarding potentially oncogenic and SFN-downregulated miRNAs, miR-29b-3p and miR-193b-3p would be interesting to study further using miRNA mimics and inhibitors, since both of these miRNAs are predicted to target a number of potential TSGs, as described in Figure 69.

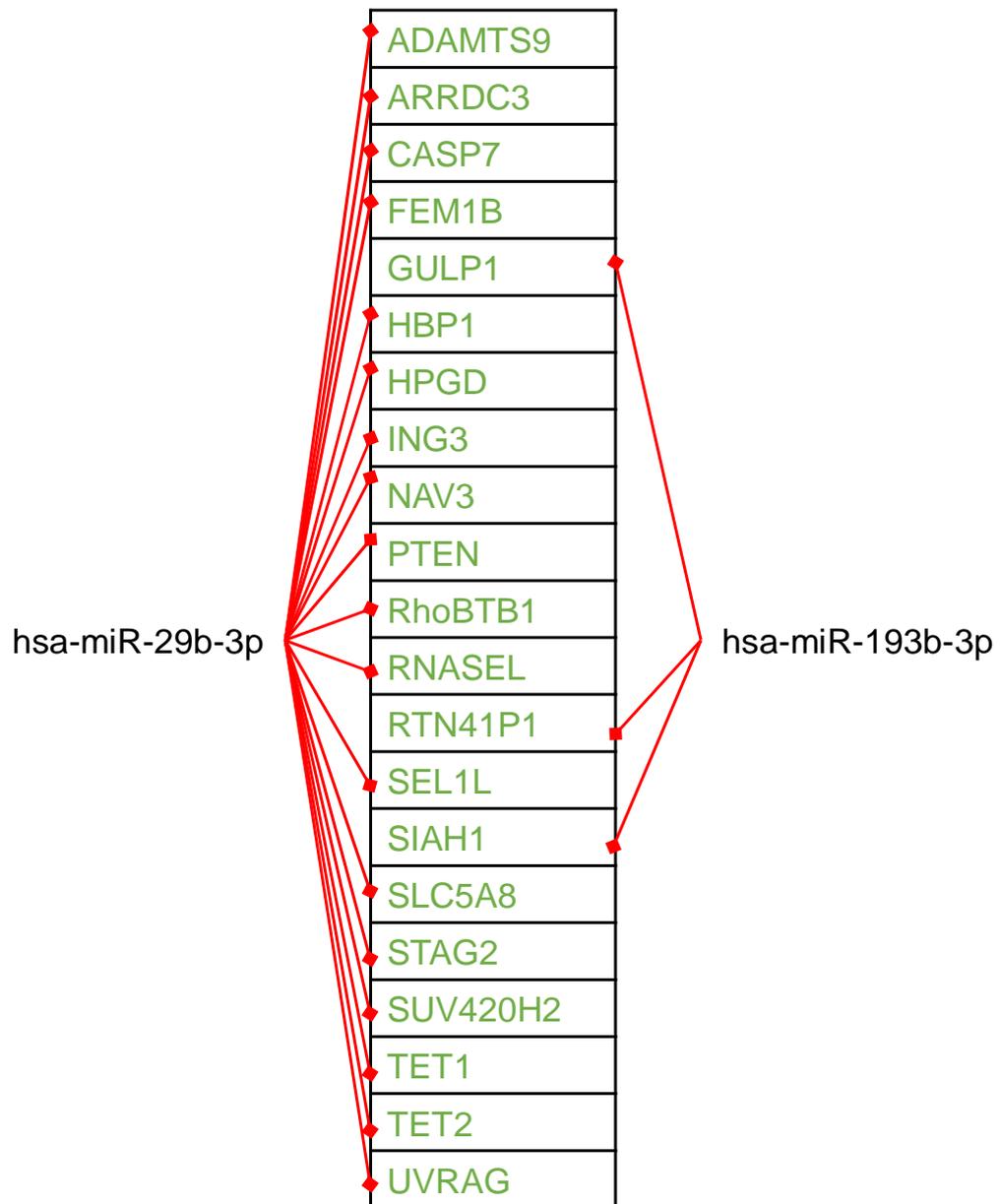


Figure 69: A diagram illustrating potentially tumour-suppressive miRNAs that are predicted by miRanda (with mirSVR scores ≤ -1.2) to be targeted by miR-29b-3p or miR-193b-3p.

6.5.1.3 Crosslinking, Ligation and Sequencing of Hybrids

In a CLASH experiment performed by Helwak et al, cells were UV-irradiated to covalently crosslink RNAs to proteins with which they were interacting, and then tagged AGO-RNA complexes were isolated. Long mRNAs in the

complexes were firstly trimmed with an exonuclease, and then fused by ligation to miRNAs present in the same complexes; this generated miRNA-mRNA 'hybrids', which were subject to deep sequencing and analysis to identify any AGO-mediated miRNA-mRNA interactions occurring *in vitro* at the time of cell irradiation (Helwak, Kudla et al. 2013).

Based on currently published protocols (Helwak, Kudla et al. 2013), individual CLASH experiments are likely to be expensive and time consuming, and thus not a viable endeavour for all researchers. However, already-published CLASH data such as those from Helwak et al. could be used to complement algorithm-based miRNA-mRNA interaction predictions, by confirming that predicted canonical miRNA-mRNA interactions actually occur and/or highlighting any non-canonical interactions that are not computationally predicted. Perhaps such CLASH data may be useful for the future development of miRNA target prediction algorithms.

Inevitably, the sensitivity of the CLASH approach depends upon miRNA and mRNA abundances, both of which vary widely according to context. Therefore, new CLASH experiments may be worth performing where economically viable. Such experiments could be done using SFN-treated and/or non-treated Caco-2 and/or CCD-841 cells; experiments in SFN-treated cells may be more sensitive to interactions involving SFN-upregulated miRNAs, while those in non-treated cells may be more sensitive to those involving SFN-downregulated miRNAs. CLASH experiments could help to uncover novel (particularly non-canonical) cancer-related miRNA-mRNA interactions, and/or reduce the amount of time wasted on studying falsely-predicted mRNA targets.

6.5.1.4 Xenograft Studies in Mice

Interestingly, *in vivo* studies have shown that let-7a-5p mimics can significantly inhibit the growth of tumour xenografts (Wang, Zhang et al. 2012), as can the experimentally-forced downregulation of Bach1 in HCT116-based xenografts; Bach1 is a validated target of some let-7 miRNA family members.

SFN itself has already been shown to inhibit tumour xenograft growth *in vivo*, although often at high doses not representative of those reasonably obtainable through diet. In a study establishing tumours in severe combined immunodeficient (SCID) mice by xenografting human CRC cell lines, supplementation of the mice with 400µmol SFN per kg body mass per day dramatically inhibited xenograft growth, reducing mean tumour weight by 70% (Chen, Tang et al. 2012).

According to United States Food and Drug Administration guidelines, approximate human dose equivalents can be extrapolated from laboratory mouse doses by the following equation (USFDA 2005):

$$\text{Human Equivalent Dose (mg/kg)} = \text{Mouse Dose (mg/kg)} \times \frac{3}{37}$$

One study reported that the amount of SFN able to be formed from extracts of 7-day-old broccoli sprouts was 5µmol SFN per gram of sprout (Guo, Yang et al. 2014). On this basis, a human dose equivalent to that used in the above-mentioned mouse xenograft study could be obtained by consuming about 390g of 7-day-old broccoli sprouts, assuming optimal SFN formation. Therefore, such doses may warrant study for their chemotherapeutic potential against moderately advanced CRCs.

It would be interesting to conduct a xenograft experiment with human CRC cell line in mice, and then supplementing mice with low-to-moderate sub-cytotoxic doses of SFN that are most likely to find use in chemoprevention or prophylaxis in high-risk and/or cured CRC patients. There is one published study in which mice were xenografted with K-ras-driven lung tumours and then treated with low-dose SFN for 3 months, showing that such supplementation did not promote tumorigenesis, despite the cytoprotective tendencies of SFN at sub-cytotoxic doses (Kombairaju, Ma et al. 2012). The daily administered doses were about 35% of those used in the study by Cheng et al. thus hypothetically obtainable from about 140g of the 7-day-old broccoli sprouts studied by Guo et al. (2014).

Regarding synergistic effects, another study tested the ability of much lower doses of SFN to act in synergy with gemcitabine to suppress the growth of human pancreatic carcinoma MIA PaCa-2-based xenografts. In this study, doses of only 16.9 μ mol SFN per kg body mass were administered to mice bearing MIA PaCa-2-based xenograft tumours, and significantly repressed tumour growth over several days in synergy with gemcitabine (Kallifatidis, Labsch et al. 2011). The theoretically equivalent human dose would be 1.37 μ mol per kg i.e. 82.2 μ mol for a human weighing 60kg – potentially obtainable from about 17g of the 7-day-old broccoli sprouts studied by Guo et al. (2014). This could be interesting regarding the potential use of SFN itself and/or dietary recommendations to boost the efficacy of ongoing chemotherapy.

6.5.1.5 Potential Clinical Applications of MicroRNA Manipulation

As mentioned above, several studies have demonstrated the ability of mimics of miRNAs, including those of the let-7 family, to both repress the growth, migration and invasion of tumour cell lines *in vitro*, and to inhibit tumour xenograft growth in mouse models. Chemotherapy that is based on miRNAs is likely to be more potent than that based on specific oncogene-targeted siRNAs, since miRNAs such as let-7f-5p are demonstrably able to downregulate a plethora of proliferative and anti-apoptotic genes. The broad-spectrum nature of such effects is particularly attractive in light of recent evidence from genomic studies to suggest that driver mutations observed in CRCs do not tend to converge with regards to which specific genes they occur in, but tend to do so with regards to which signalling pathways are impacted; such convergence is often on the Wnt, TGF β , PI3K, receptor tyrosine kinase-RAS and p53 signalling pathways (Yamagishi, Kuroda et al. 2016).

Challenges presented by the clinical application of miRNAs include ensuring their stability in the body and efficient delivery to target cells, in addition to concerns over potential side effects of the nucleic acid delivery vehicles, and/or off-target effects of the mimic itself. However, research in recent years has focussed on overcoming such challenges, to be able to develop potentially safe and effective miRNA-based drugs.

MRX34 was developed by Mirna Therapeutics (Texas, USA) and consists of natural miR-34a-5p – a miRNA reported to target a number of oncogenes - encapsulated by a liposomal nanoparticle. However, a phase I clinical trial of injection-based MRX34 chemotherapy was recently terminated due to the occurrence of serious immune-related adverse events (ClinicalTrials.gov: NCT01829971). An alternative to liposomal delivery is to package miRNA mimics within bacterium-based nano-cells, such as the EDV™ Nanocells (EDVs) developed by EnGeneIC, and then coat the EDVs with antibodies raised against antigens known to be abundant on the target cancer cell plasma membrane (Reid, Kao et al. 2016). A phase I clinical safety trial of a drug based on EDVs containing mimics based on the miR-15/107 consensus sequence, coated with EGFR-targeting antibody, is currently recruiting participants with malignant pleural mesothelioma or non-small cell lung cancer (ClinicalTrials.gov: NCT02369198).

An alternative method of upregulating miRNA-mediated repression is to target endogenous repressors of miRNA activity, such as the lncRNA H19, which has been demonstrated to act oncogenically in CRC. H19 has been shown to act as a decoy for let-7 miRNAs, amongst other miRNAs, thereby inhibiting let-7-mediated repression of oncogenic mRNA translation (Liang, Fu et al. 2015). Interestingly, another lncRNA, ncNRFR, also appears to inhibit let-7-mediated gene repression, but by competitively binding to the mRNA let-7 target sites, rather than by binding to the miRNAs themselves (Franklin, Rankin et al. 2013).

The approach of targeting such lncRNAs might be more effective than the administration of let-7 miRNA mimics in elevating let-7-mediated oncogene repression, particularly in contexts of high endogenous let-7 miRNA levels. Caco-2 cells appear to have high basal let-7f-5p and let-7g-5p expression, as implied by relatively high normalised library read numbers and strong Northern Blot signals for these miRNAs in Caco-2 samples. That they were significantly upregulated by SFN treatment in terms of fold-change implies a steep increase in absolute miRNA levels. High endogenous let-7f-5p levels in Caco-2 were the most probable cause of the difficulty in detecting effects of let-7f-5p mimic via luciferase assays in those cells, even when vectors

containing 'positive-control' repeats of sequence anti-sense to that of let-7f-5p were used. The targeting of lncRNAs such as H19 by siRNA-based knockdown is a potential alternative and/or complementary means of increasing let-7-mediated oncogene repression in CRC.

6.5.1.6 Clinical Applications of Sulforaphane

A synthetic SFN supplement called Sulforadex®, developed by Evgen Pharma, has now entered phase II safety trials for the treatment of metastatic breast cancer patients in combination with aromatase inhibitors, tamoxifen and fulvestrant (ClinicalTrials.gov: NCT02970682).

With regards to the efficacy of SFN and/or SFN-rich extracts in the treatment of cancer patients, a pilot study is currently recruiting participants to determine the feasibility of a randomised controlled trial treating advanced pancreatic ductal adenocarcinoma patients – receiving palliative treatment - with freeze-dried broccoli sprouts (ClinicalTrials.gov: NCT01879878).

It would be interesting to follow such clinical trials, particularly to see whether broccoli sprout and/or SFN supplementation prove to be beneficial, neutral, or detrimental to patient outcomes. Since SFN induces Nrf2, whose cytoprotective effects can inhibit the carcinogenic transformation of healthy cells, yet can promote tumour growth when occurring in tumour cells (Ohta, Iijima et al. 2008), concerns have been raised over the possibility of low-to-moderate doses of SFN being detrimental to patients with advanced cancers via tumour protection, as previously discussed. However, no clinical studies published to date have actually demonstrated such detrimental effects in practice. Besides, SFN is known to affect a wide range of Nrf2-independent signalling pathways, such as those modulating cell cycle arrest, apoptosis and chromatin status, as previously discussed. Such effects may outweigh the oncogenic potential of Nrf2 induction in tumour cells, and it is important to remember that Nrf2 induction in tumour cells can itself be tumour suppressive rather than oncogenic, as discussed in Chapter 1. Additionally, the systemic effects of Nrf2 activators will inevitably extend far beyond those observable *in vitro* due to potential impacts upon immune system-mediated tumour suppression and/or promotion.

Appendix

7.1 Additions to Chapter 1

7.1.1 MicroRNA Nomenclature

7.1.1.1 Introduction of the -5p/-3p Suffix

As mentioned in the MicroRNAs section in Chapter 1, it was once considered that only one strand of a Dicer-generated duplex tended to exert activity (Mah, Buske et al. 2010). The other ‘alternative’ strand was thus denoted by a “*” suffix in previous versions of miRBase. For example, of the two mature products of pre-hsa-miR-29b-1/pre-hsa-miR-29b-2 cleavage, the ‘dominant’ product from the 3’ end was named “hsa-miR-29b”, whilst the ‘alternative’ product from the 5’ end was named “hsa-miR-29b*”.

As accumulating evidence continued to challenge the notion that this was the case for the majority of miRNAs, this dominant/alternative strand notation was discontinued by miRBase; instead, a -5p or -3p suffix is appended to indicate the arm of the pre-miRNA from which the mature miRNA strand originates. For example:

Old miRBase Nomenclature	New miRBase Nomenclature
hsa-miR-29b	hsa-miR-29b-3p
hsa-miR-29b*	hsa-miR-29b-5p

According to this current nomenclature, the lack of a -5p or -3p suffix on a named mature miRNA in miRBase implies that only a single mature product has been thus far reported.

7.1.1.2 IsomiRs

Currently in miRBase, only a single consensus (canonical) sequence is listed for any named mature miRNA, such as “hsa-miR-29b-3p”. MiRBase does not currently employ nomenclature to define different isomiRs.

Canonical isoforms (as catalogued in miRBase) were here referred to according to standard miRBase nomenclature, while to non-canonical isomiRs were appended extensions according to a system proposed by Cloonan et al. (Cloonan, Wani et al. 2011); “[isomiR]_{x_y}”, where *x* is the position of the pre-miRNA at which the miRNA isoform begins, and *y* is the position of the pre-miRNA at which it ends, where the first nucleotide of the pre-miRNA is at position 0.

For example:

	Isoform	Name Used Here
hsa-miR-29b-3p	Canonical – spans positions 51 to 73 according to miRBase (first position = 1).	hsa-miR-29b-3p
	IsomiR – misses the first nucleotide, and has 2 extra nucleotides at the end from the precursor; spans positions 52 to 75 according to miRBase (where the first position is defined as position 1).	hsa-miR-29b-3p[isomiR]51_74

If an isomiR has any base ‘substitutions’ compared with the canonical sequence, these are denoted with a further extension: “sub.*P.A>B*”, where *P* is the position of substitution in the pre-miRNA (where the first position is defined as position 0), and *A* and *B* are the bases found in the canonical and the described isomiR respectively. For example, a hsa-miR-29b-3p isomiR that differs from the canonical form by missing the first nucleotide, having 2 extra nucleotides at the end from the precursor, and having a substitution of A for G at position 61 (where the first position is defined as position 0) with respect to the pre-miRNA, would be described as:

“hsa-miR-29b-3p[isomiR]51_74|sub.61.G>A”

7.2 Additions to Chapter 4

7.2.1 MicroRNA Library Construction

7.2.1.1 Checking RNA Integrity

The integrity of extracted total RNA samples was checked prior to library construction, to make sure that RNA was not degraded. When denatured and then subject to agarose gel electrophoresis, intact RNA is expected to show clearly defined 28S and 18S bands, with the intensity of the latter being about half that of the former.

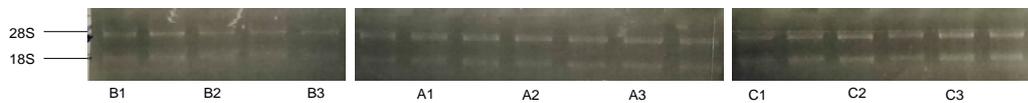
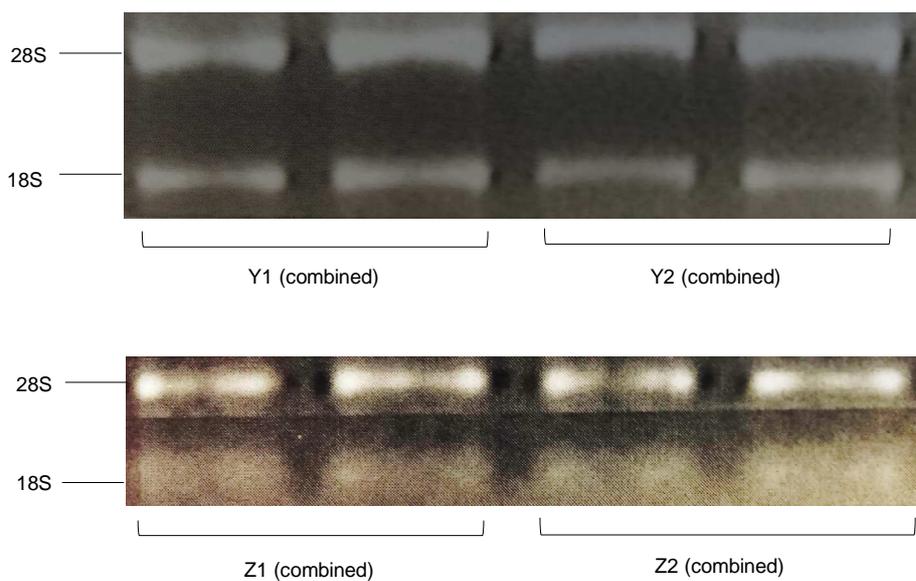


Figure 70: Ethidium bromide-stained images of the agarose gel through which Caco-2 total RNA samples were electrophoresed, in order to check RNA integrity prior to library construction. The samples subsequently used for library construction are annotated below the gel image.

Figure 70 shows the electrophoresis images for all of the RNA samples used for the construction of the Caco-2 libraries, while Figure 71 shows those from CCD-841 cells used for CCD-841 library construction.



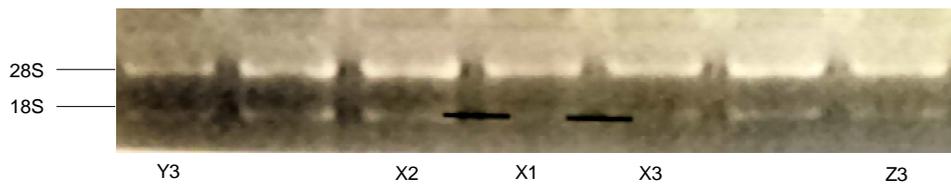


Figure 71: Ethidium bromide-stained image of the agarose gel through which CCD-841 total RNA samples were electrophoresed, in order to check RNA integrity prior to library construction. The samples subsequently used for library construction are annotated below the gel image.

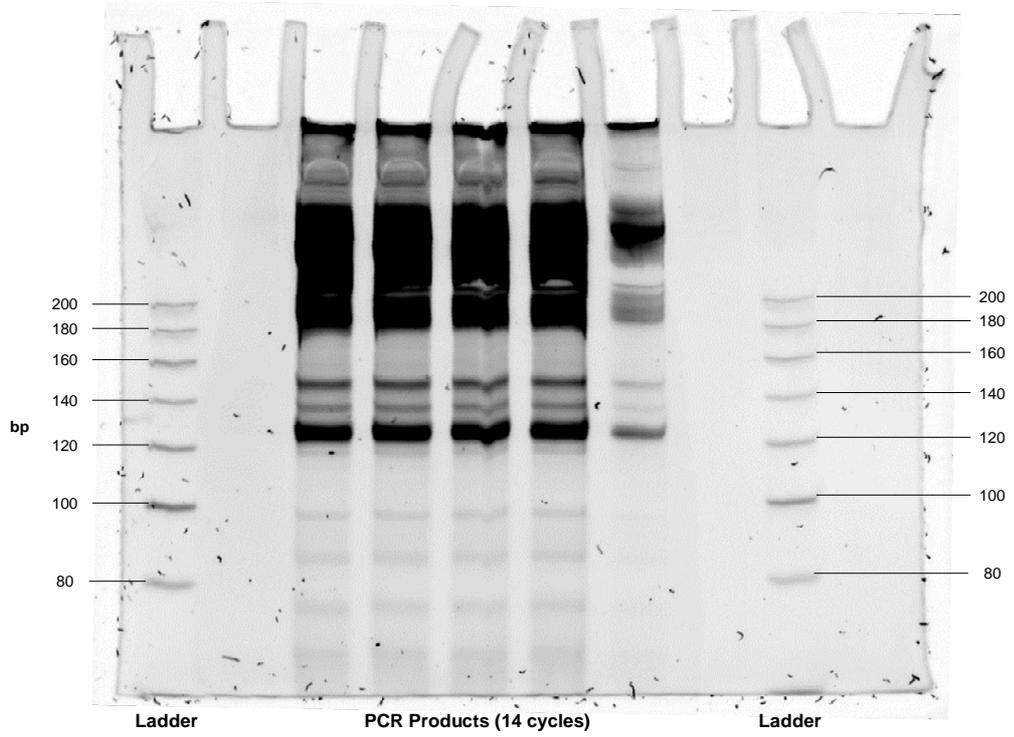
It is clear from the images in Figure 70 and Figure 71 that no RNA degradation is apparent for any of the samples, since all lanes have clearly defined 28S and 18S bands, and the intensities of the 18S bands are lower than those of their respective 28S bands. Thus, these samples were considered to be of suitable quality for library construction.

7.2.1.2 Elution of the Libraries

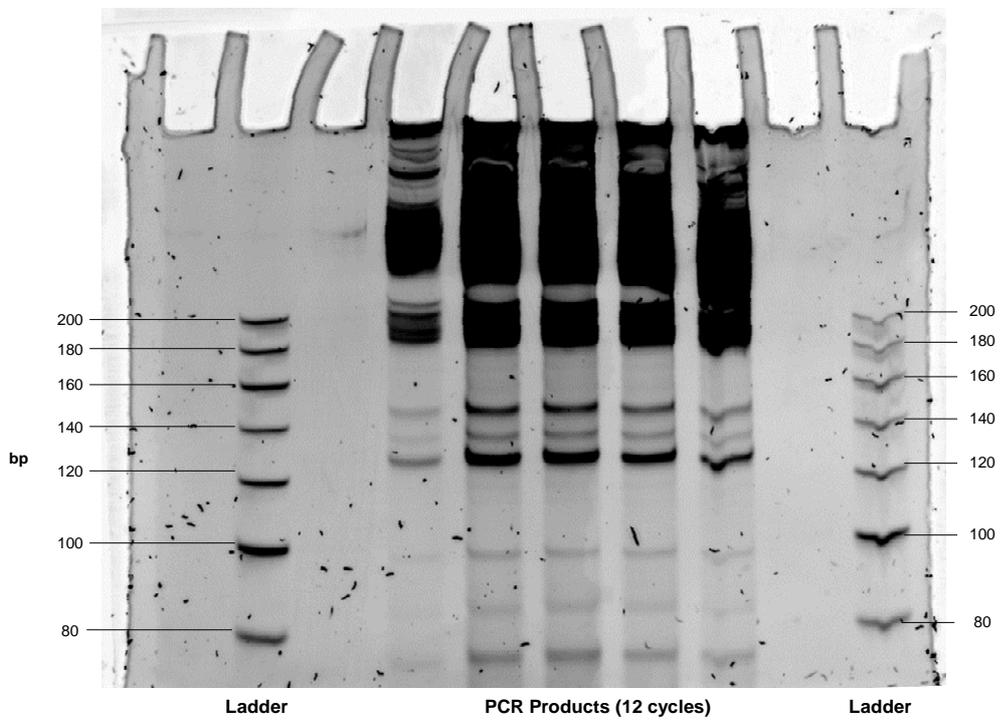
Following the ligation of 5' and 3' HD adapters to small RNAs, reverse transcription and subsequent PCR amplification, the PCR products were separated by PAGE. The desired microRNA cloning products were to be found in the band located at around 150bp, while the adapter-adapter dimers formed bands below 140bp and were avoided. Pieces of the gel at the loci of the desired (150bp, microRNA) bands were carefully cut out, and the DNA eluted from the gel to form the library.

Different numbers of PCR cycles to test which number would produce the most of the desired products without the adapter-adapter dimer bands diffusing over and contaminating the microRNA bands. If there was a clear gap between all of the microRNA and adapter-adapter dimer bands then all of the desired bands were cut out for elution. However, if the PCR cycle number appeared to be too high in any lanes, then PCR of the original cDNA was performed again using the 'optimal' number of cycles, and another PAGE separation was performed for elution. Images of all of the PAGE gels from which the microRNAs libraries that were sent for deep sequencing were cut out and eluted are shown in Figure 72.

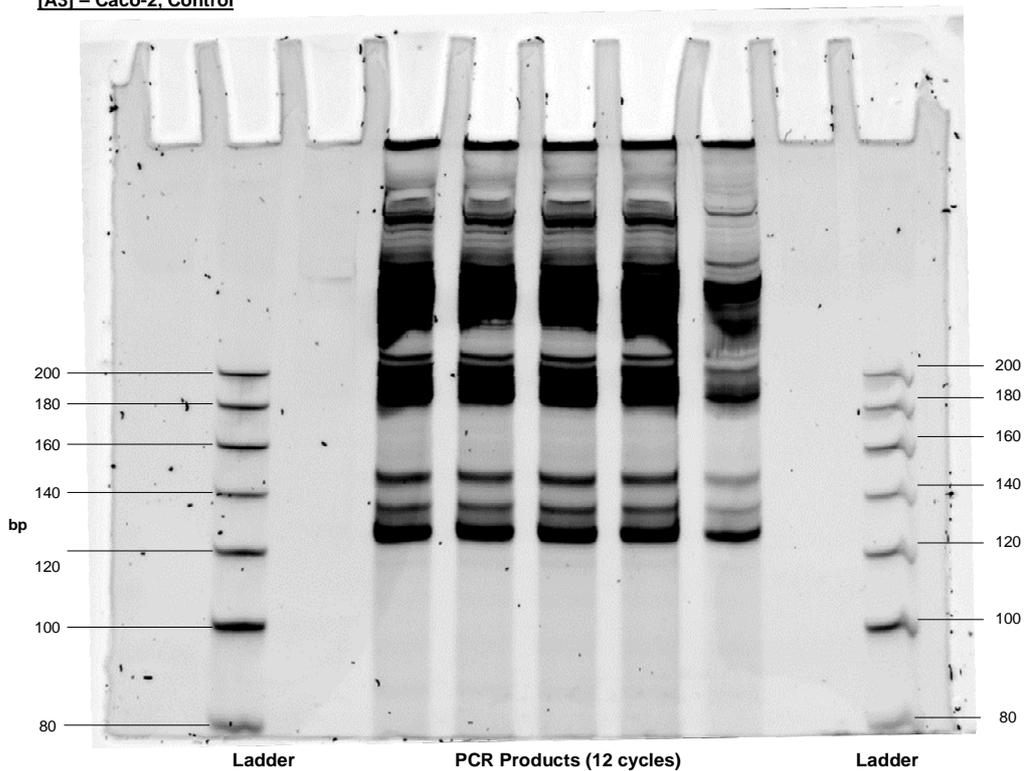
[A1] – Caco-2, Control



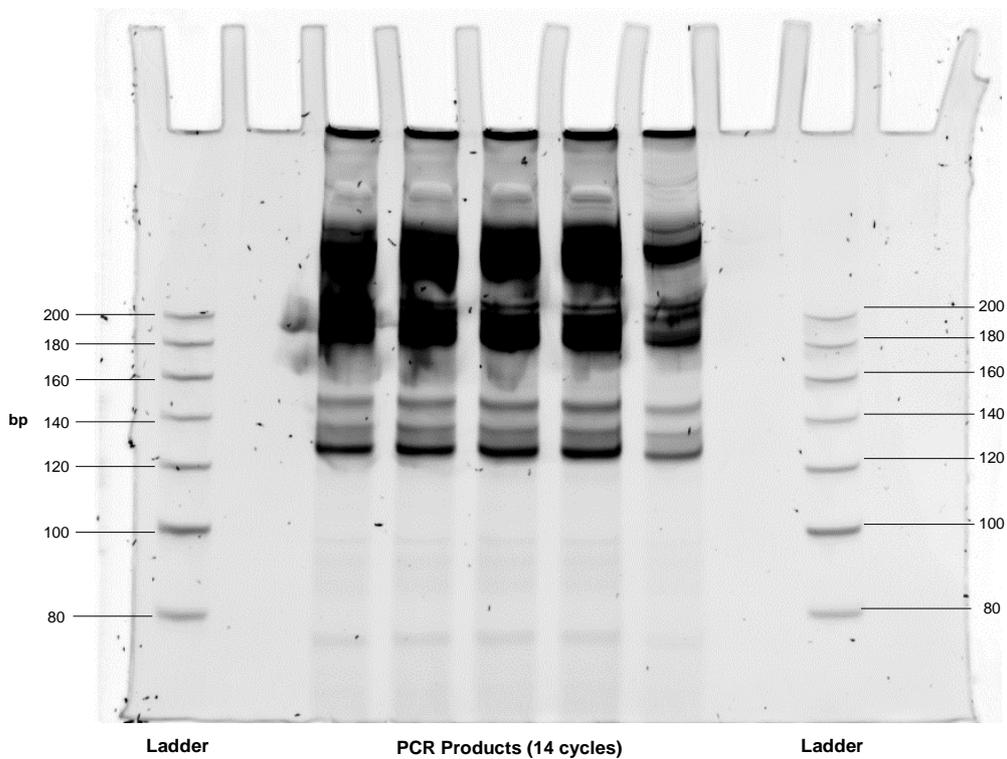
[A2] – Caco-2, Control



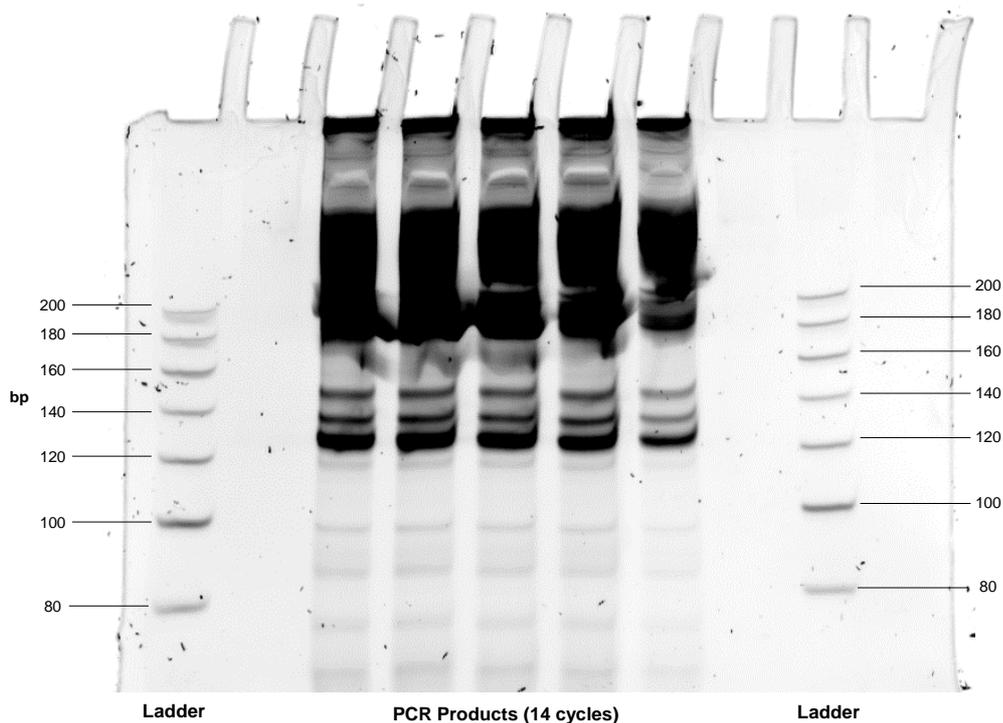
[A3] – Caco-2, Control



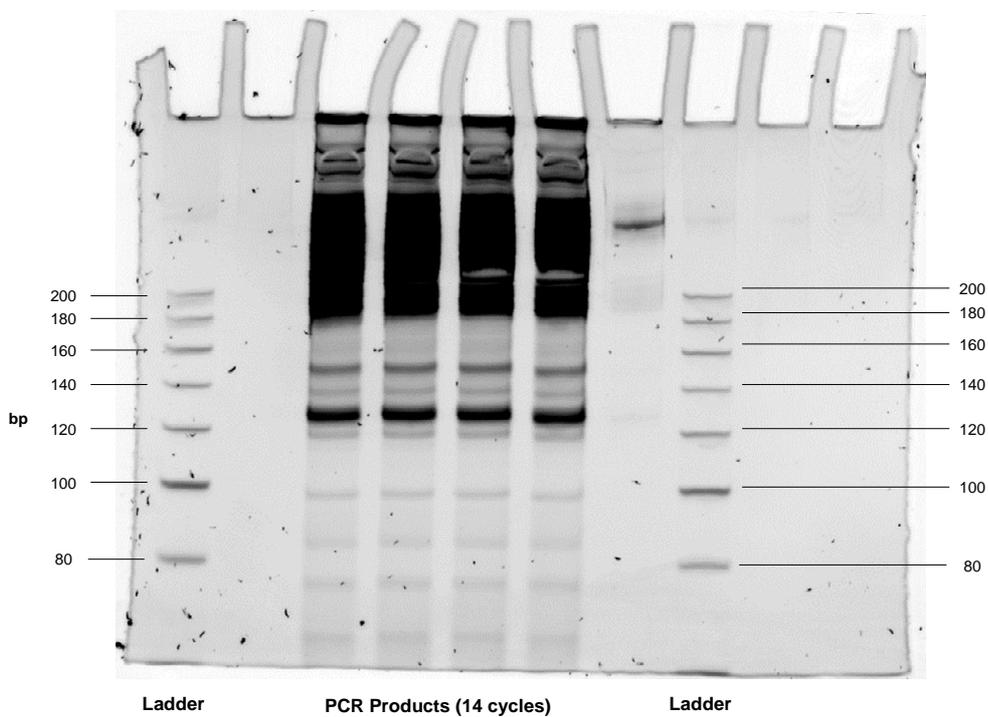
[B1] – Caco-2, 8 h SFN



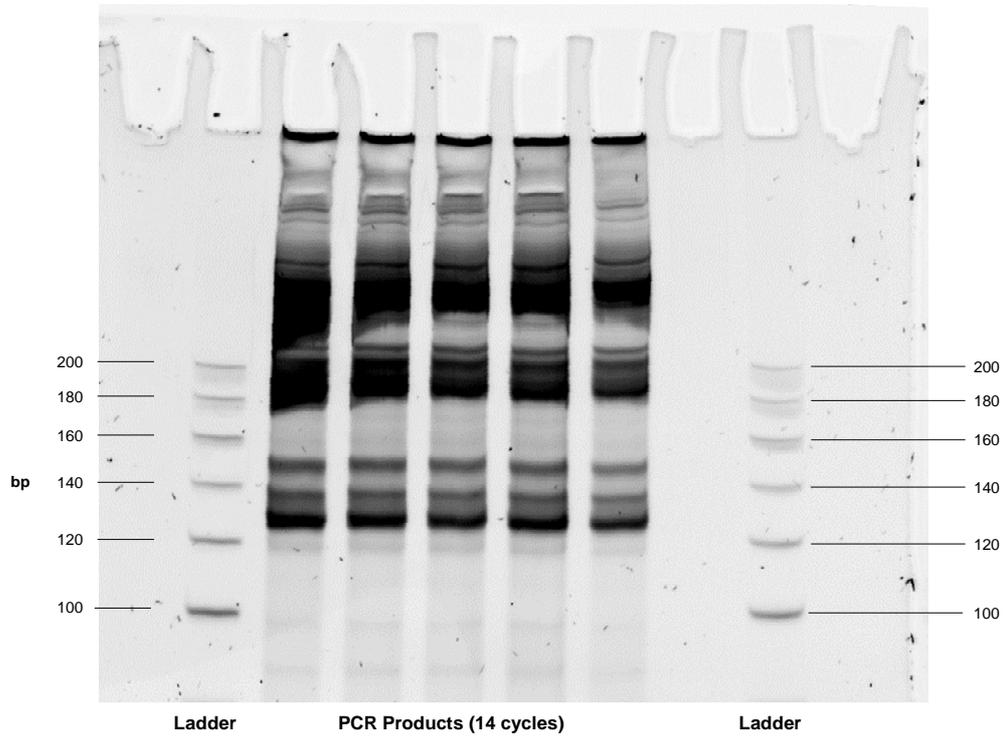
[B2 – Caco-2, 8 h SFN



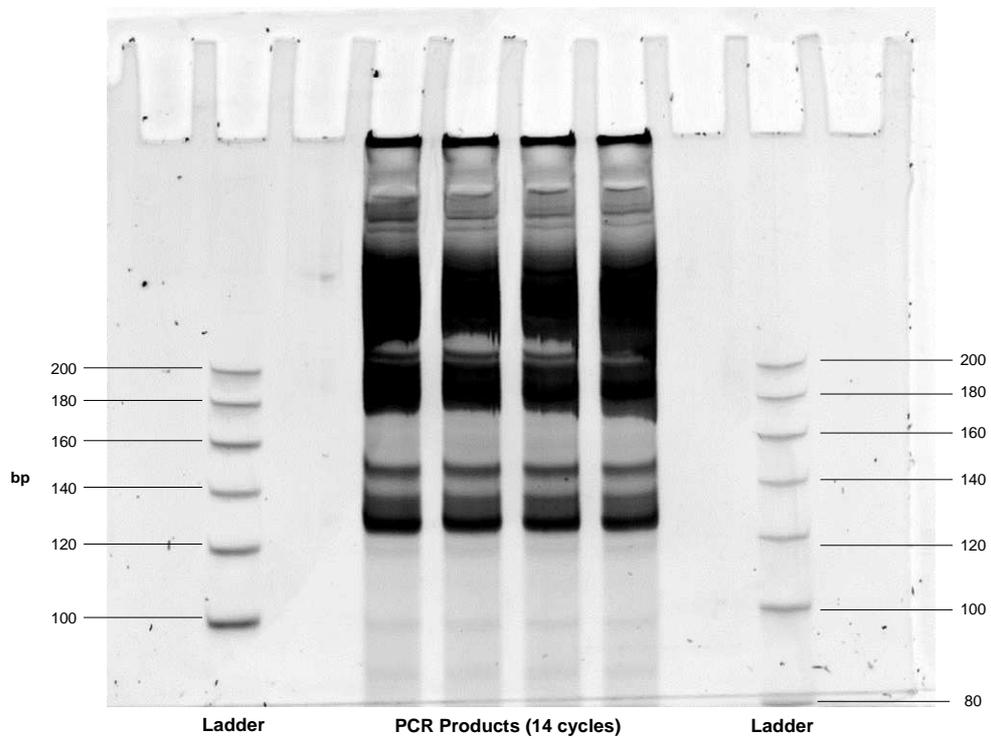
[B3 – Caco-2, 8 h SFN



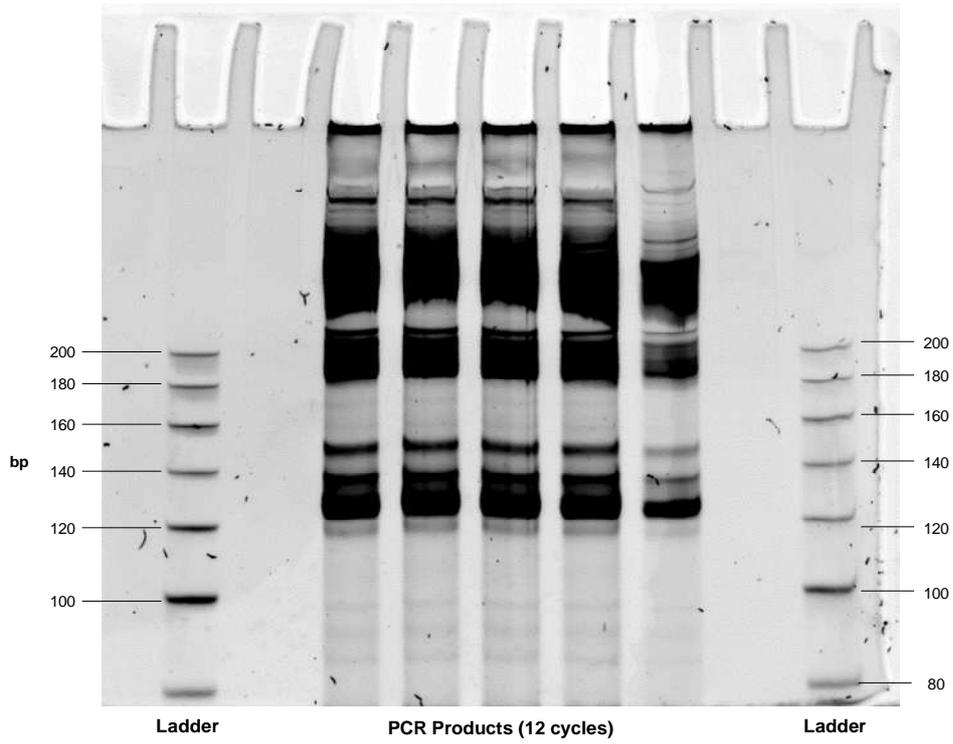
[C1 – Caco-2, 24 h SFN]



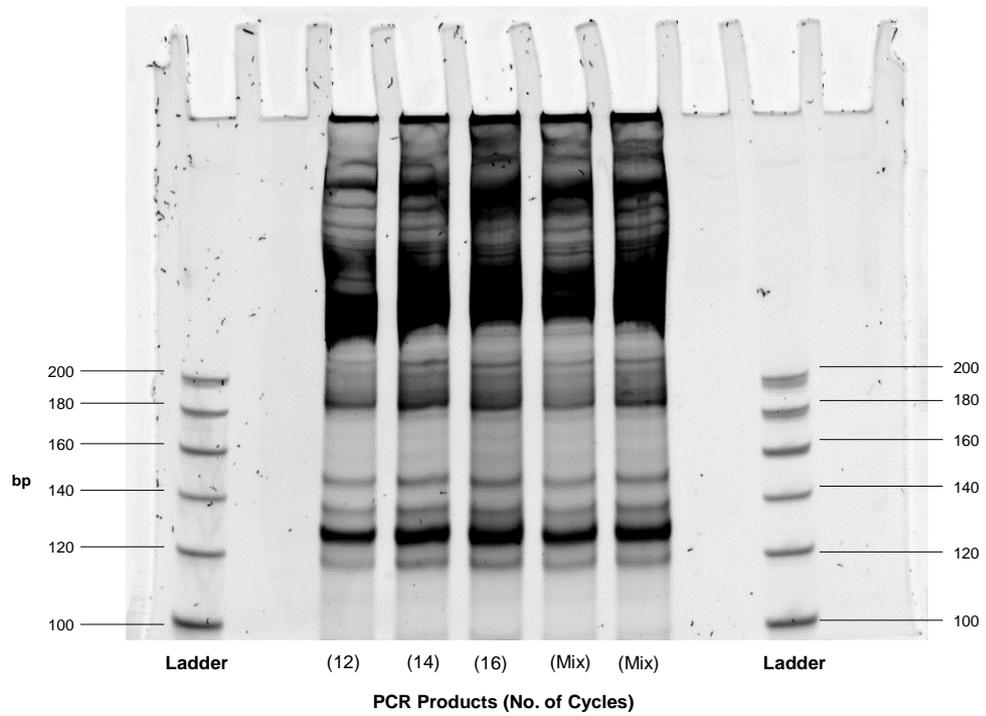
[C2 – Caco-2, 24 h SFN]



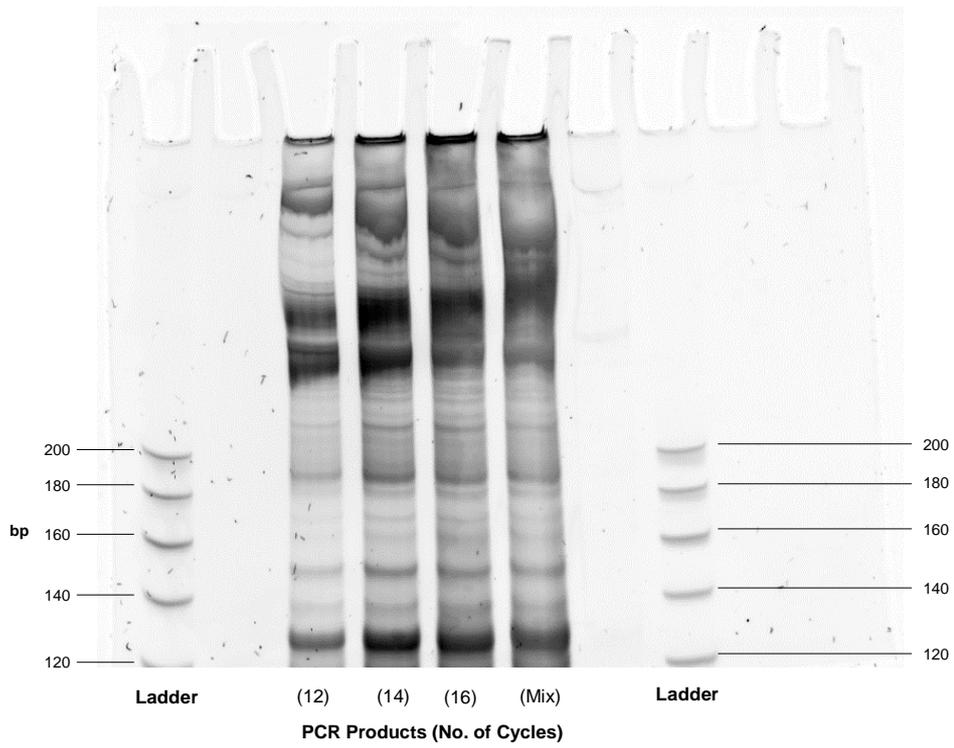
[C3 – Caco-2, 24 h SFN]



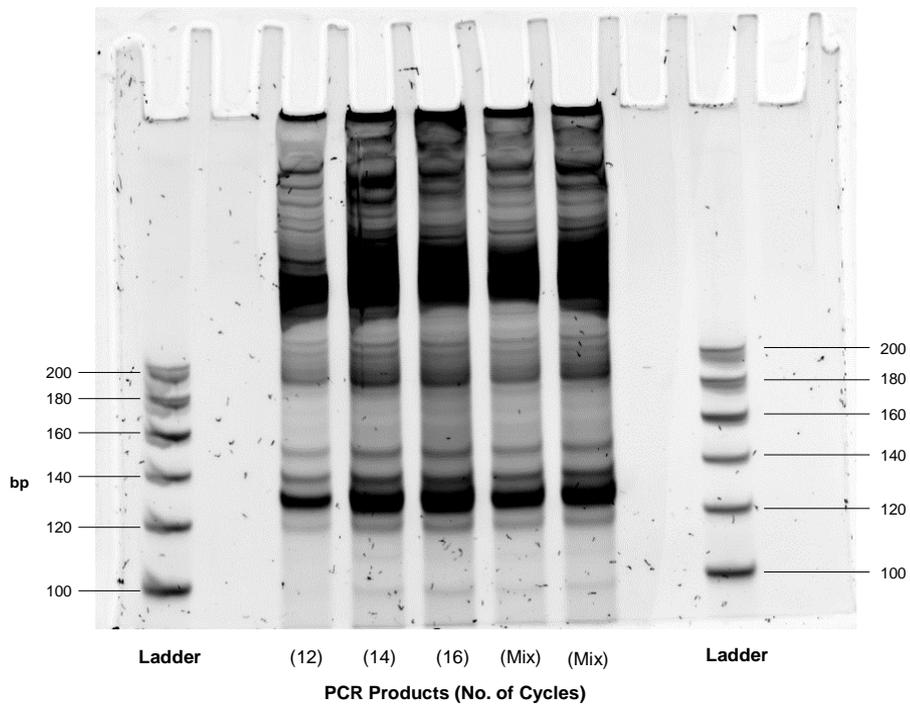
[X1] – CCD-841, Control



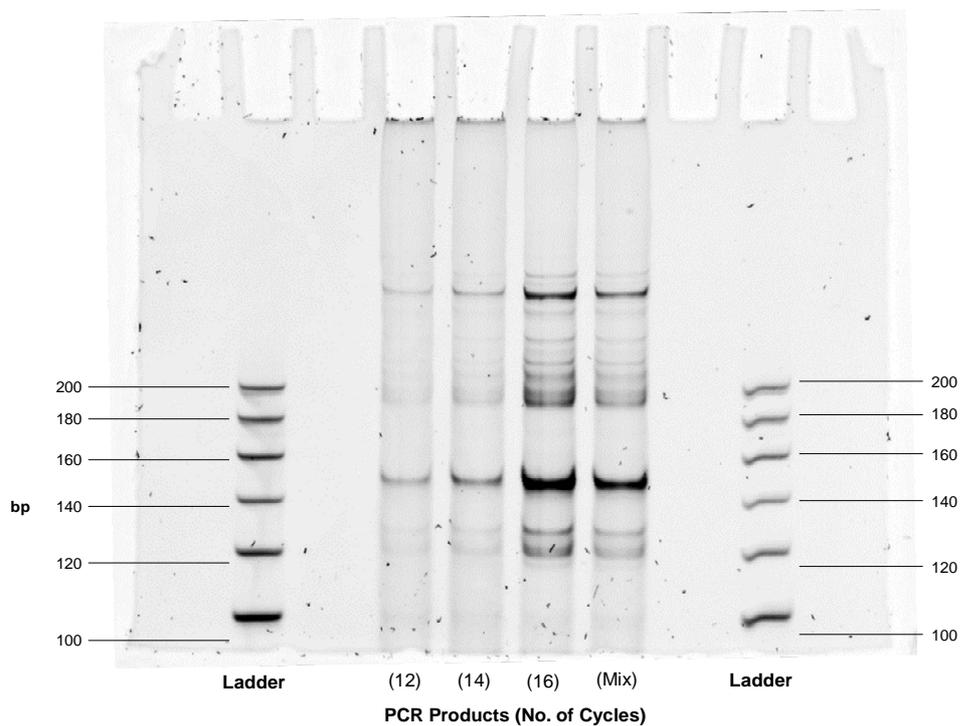
[X2] – CCD-841, Control



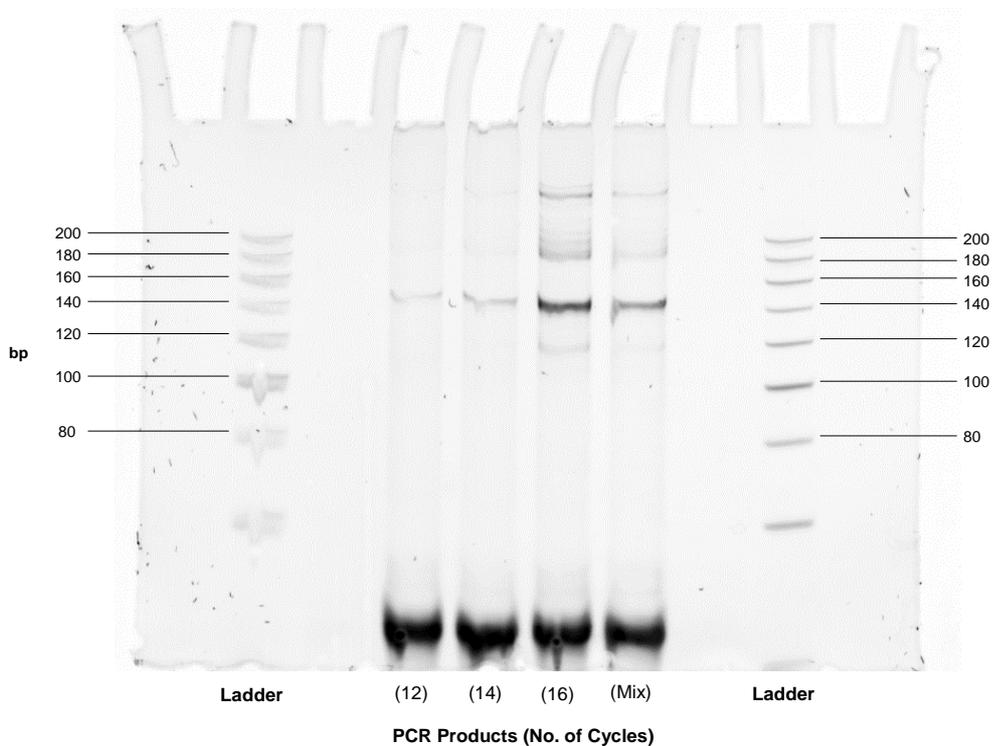
[X3] – CCD-841, Control



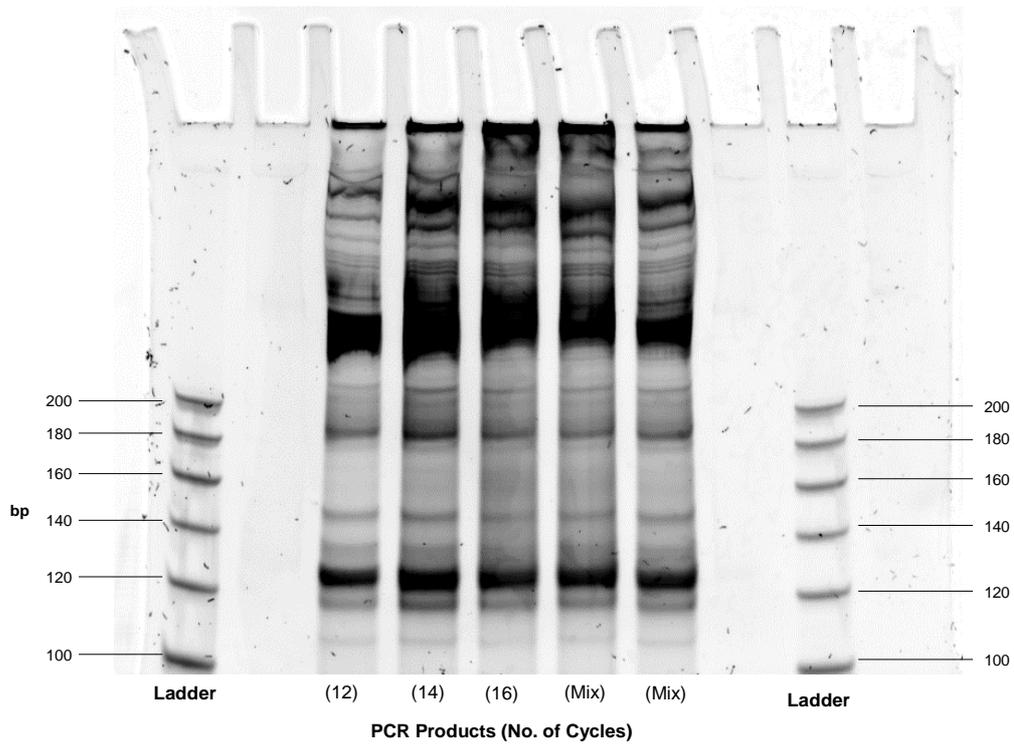
[Y1] – CCD-841, 8 h SFN



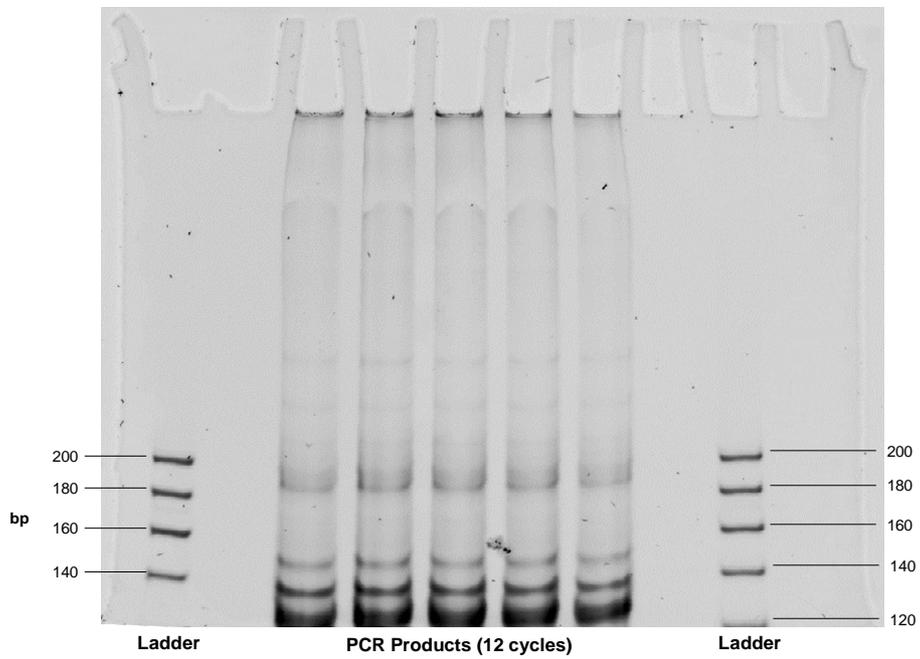
[Y2] – CCD-841, 8 h SFN



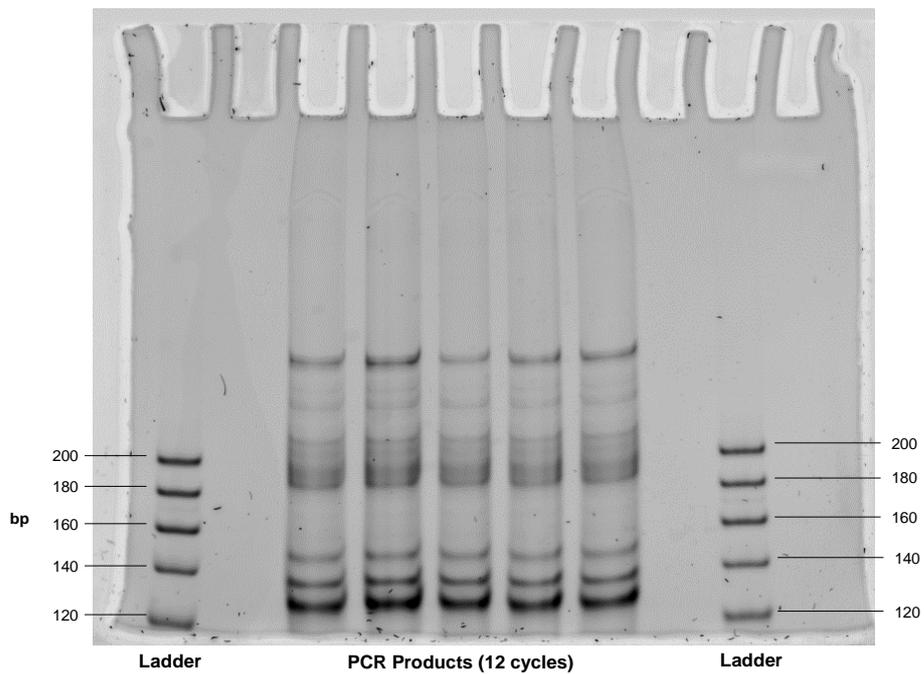
[Y3] – CCD-841, 8 h SFN



[Z1] – CCD-841, 24 h SFN



[Z2] – CCD-841, 24 h SFN



[Z3] – CCD-841, 24 h SFN

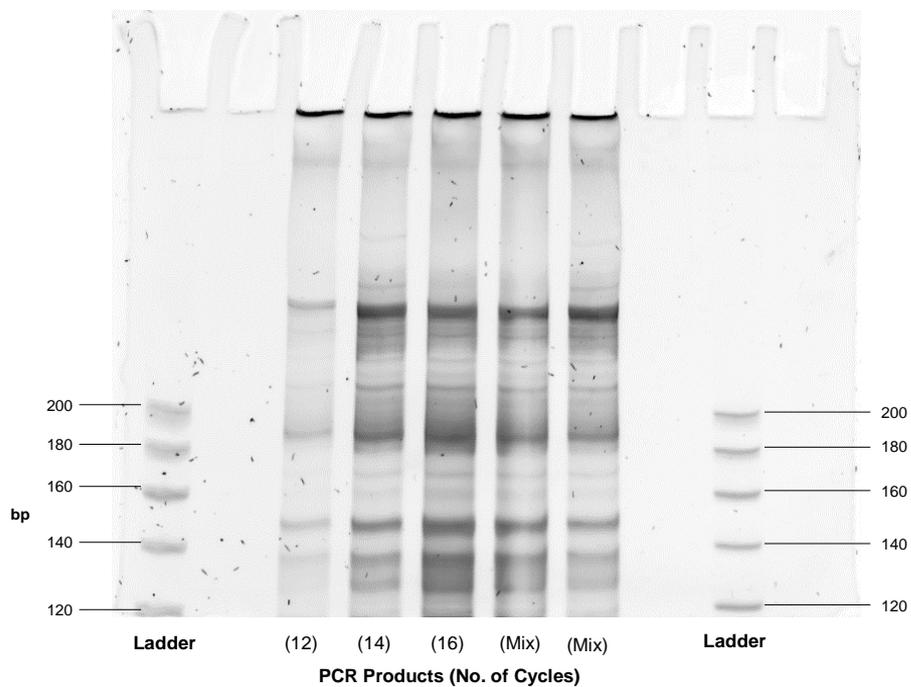


Figure 72: Images of ethidium bromide-stained gels with PAGE-separated PCR products of cDNA that was reverse transcribed from small RNAs captured from samples by the 5' and 3' HD adapters. The bands visible between 140 and 160bp are the desired microRNA bands, which were cut out and from which libraries were eluted, while those visible below are from the adapter-adapter dimers, which care was taken to avoid.

7.2.1.3 Quality Check and Quantification

All Caco-2 libraries were combined into one pool, while those of CCD-841 were combined into another pool. Prior to mixing libraries together into pools, aliquots of each were separated by PAGE to both check the specificity of the library gel extraction (i.e. that other products were not extracted at the same time, particularly AADs), and to quantify the relative concentration of each library preparation, so that equal amounts of each could be pooled together.

7.2.1.3.1 Caco-2 Libraries

1 μ L of each library preparation was electrophoresed in a PAGE gel and then bands were quantified to estimate relative concentrations. The data are shown in Figure 73.

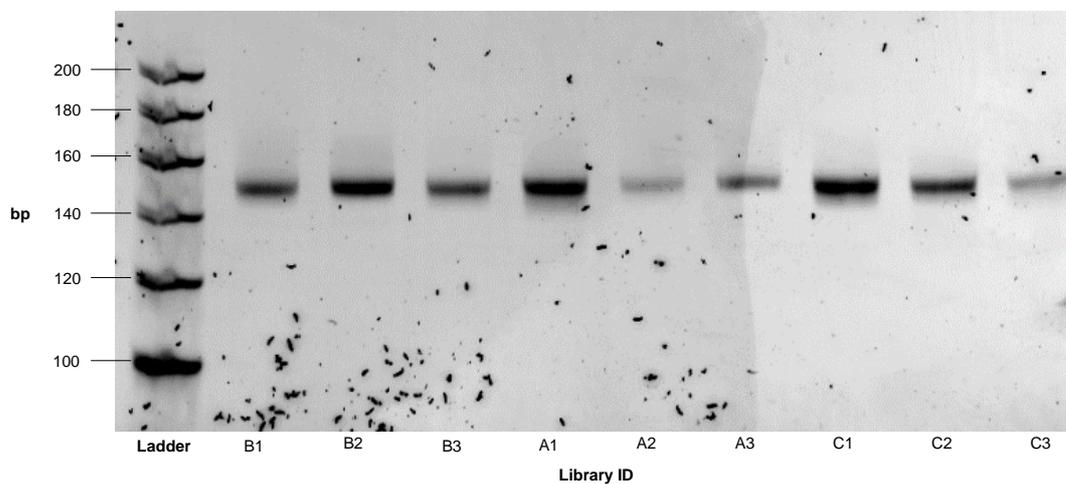


Figure 73: Ethidium bromide-stained gel with Caco-2 libraries electrophoresed for quality checking and quantification of concentrations.

As apparent from the image in Figure 73, there does not appear to be contamination of the eluates with significant amounts of AADs or other undesired products. Band intensities were densitometrically analysed to estimate relative library concentrations, based on which, equal amounts of each library were then combined into a pool to be sent for deep sequencing.

7.2.1.3.2 CCD-841 Libraries

1 μ L of each library preparation was electrophoresed by PAGE across two gels, alongside a reference product that was loaded onto both gels in order to normalise results across the gels. Data are shown in Figure 74 and Figure 75.

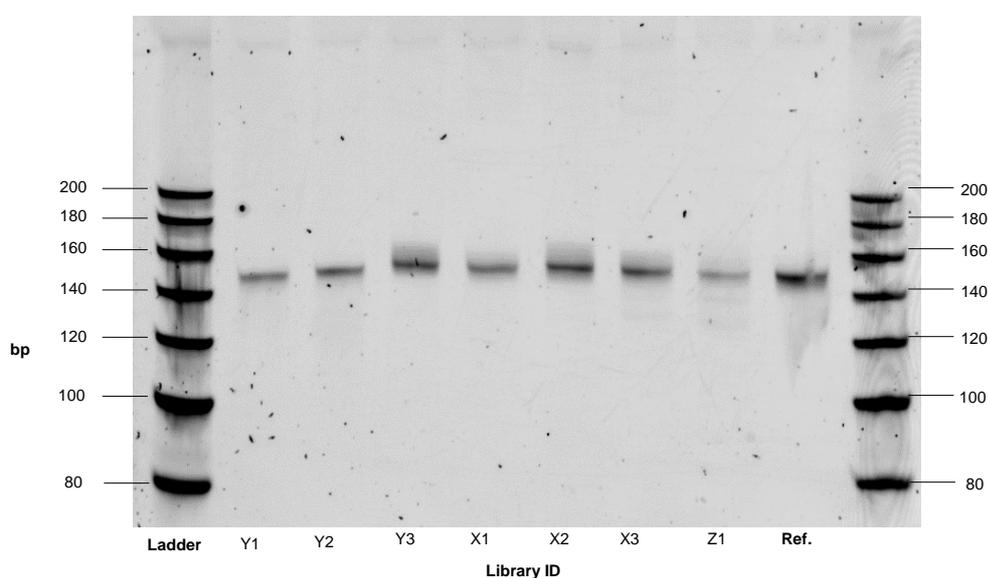


Figure 74: The first of two ethidium bromide-stained gels with CCD-841 libraries electrophoresed for quality checking and quantification of concentrations.

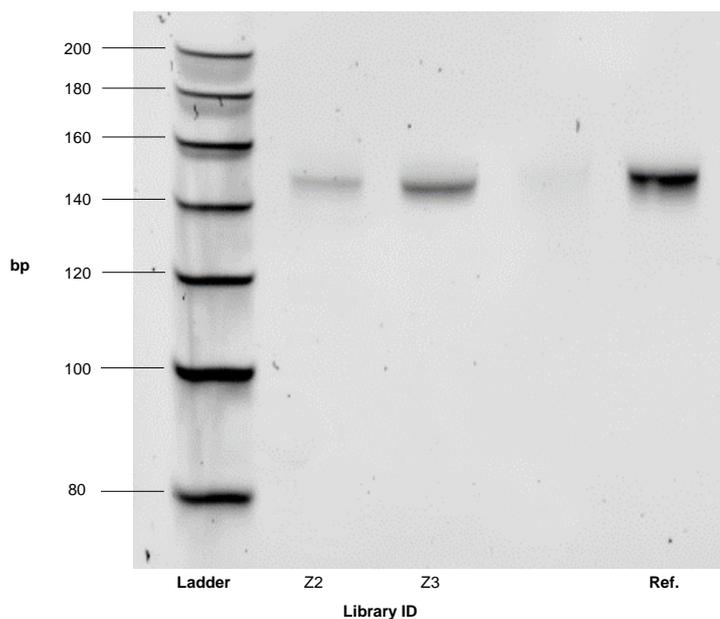


Figure 75: The second of two ethidium bromide-stained gels with CCD-841 libraries electrophoresed for quality checking and quantification of concentrations.

The images in Figure 74 and Figure 75 indicate that none of the CCD-841 libraries have significant contamination with undesired products of differing length, such as adapter-adapter dimers. Band intensities were densitometrically analysed to estimate relative library concentrations, based on which, equal amounts of each library were then combined into a pool to be sent for deep sequencing.

7.2.2 MicroRNA Library Deep Sequencing Data Analysis

7.2.2.1 Size Class Distributions

The size class distribution data presented here were generated by Claudia Paicu (School of Computing Sciences, University of East Anglia, Norwich, UK).

7.2.2.1.1 Caco-2 Libraries

Table 12: A summary of the size class distributions of reads from the sequencing of the Caco-2 libraries.

	A1	A2	A3	B1	B2	B3	C1	C2	C3
17	624415	493750	57600	314086	106181	303613	679836	607286	885765
18	702375	709236	119727	442857	194891	480731	811991	809983	976872
19	681733	806697	271418	562627	312534	517538	951718	1003180	1067874
20	1108053	1429021	758037	931522	637199	803716	1537759	1623031	1688904
21	2300053	2968175	2313316	1980115	1532961	1469488	3155053	2979848	3425674
22	5086505	6830196	4502954	3669751	3914985	3574137	5387507	5417295	6052942
23	4098896	5942047	4128015	3381937	3541925	2982853	4773403	4352262	4862898
24	1493362	2797371	1903376	1687144	1612057	1410744	2153280	1930011	1949966
25	255787	695766	450264	414743	374633	353870	467457	418809	356968
26	92174	355808	215215	216962	174323	157146	180850	149924	123869
27	43841	278925	166925	176860	124388	84574	102241	77921	73807
28	22546	226934	147281	153529	98718	43733	58625	41158	39077
29	11028	149533	94516	112945	76652	21558	31058	22416	17929
30	5942	93150	44172	71686	51893	12808	15062	11252	8437
31	4535	83132	28601	72839	57296	9942	13312	9403	7180
32	3404	57704	13664	63435	57555	7662	8135	5679	4966

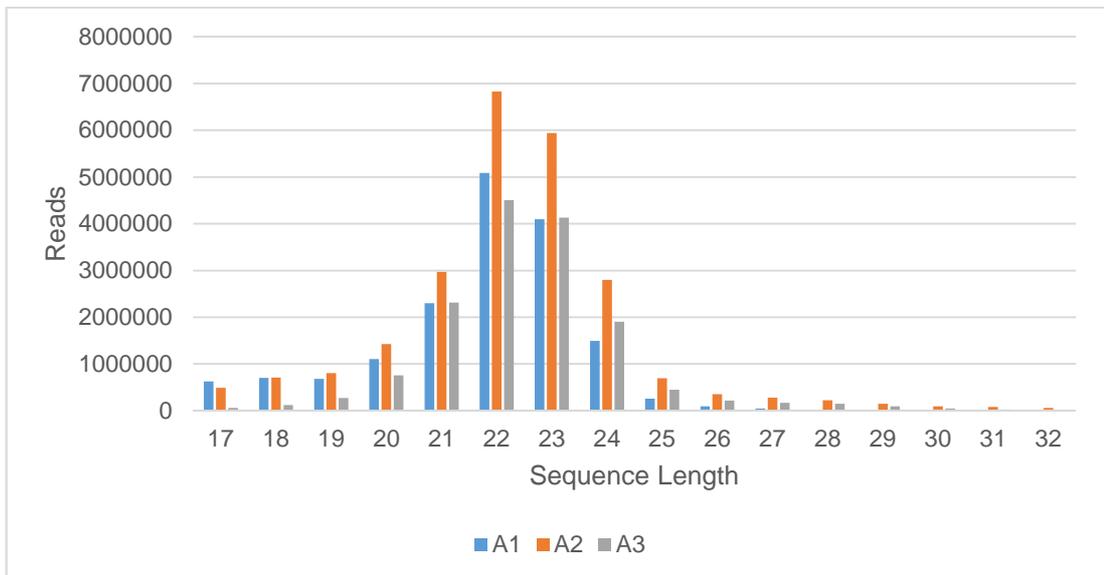


Figure 76: A graph representing the Caco-2 size class distribution data for A1, A2 and A3.

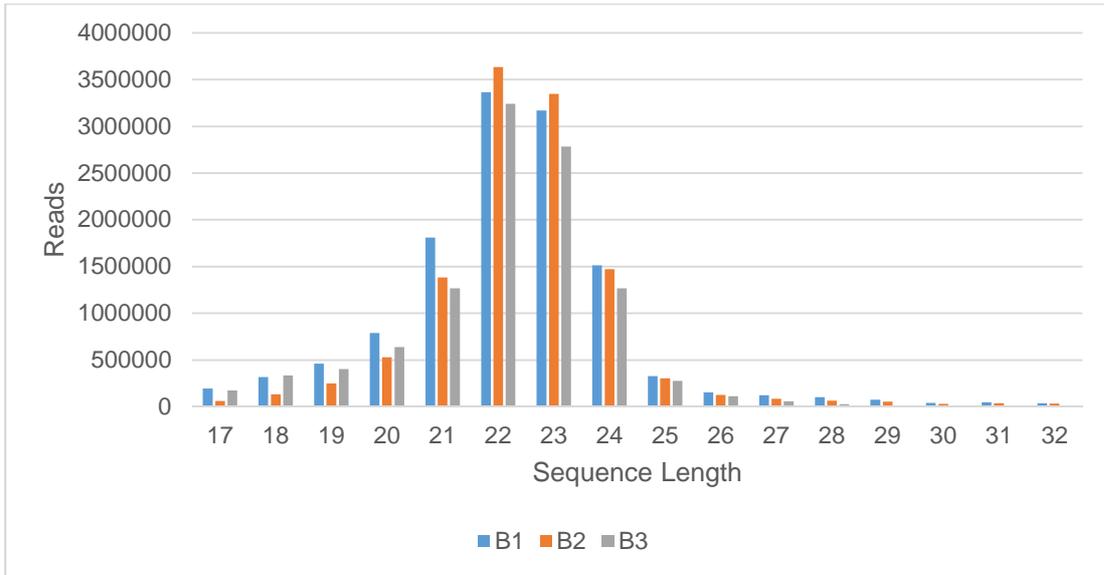


Figure 77: A graph representing the Caco-2 size class distribution data for B1, B2 and B3.

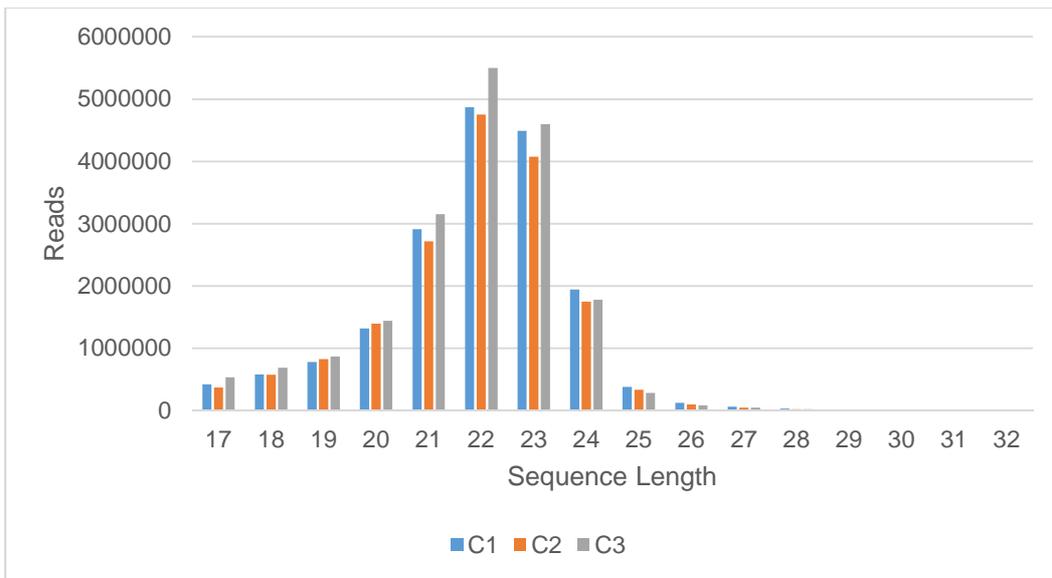


Figure 78: A graph representing the Caco-2 size class distribution data for C1, C2 and C3.

The size class distributions for all of the Caco-2 libraries are as expected for human miRNAs; peaks at 22 nucleotides are clearly visible in all of them as visible in Figure 76, Figure 77 and Figure 78.

7.2.2.1.2 CCD-841 Libraries

Table 13: A summary of the size class distributions of reads from the sequencing of the CCD-841 libraries.

	X1	X2	X3	Y1	Y2	Y3	Z1	Z2	Z3
17	262327	169940	121180	734821	786455	95181	353461	172476	100725
18	373405	246985	178375	1488664	1596769	203134	384207	236563	249435
19	407730	278480	236282	3971359	4318924	303211	523853	309847	480325
20	701314	336343	416106	5465803	5630577	514593	843315	604590	1053300
21	1898162	764814	1143511	11537303	12283235	1234124	2214310	1658868	3262699
22	7153520	3643959	4400003	885721	595400	5207008	6911060	5160677	16322852
23	3114046	1174147	1994755	390896	260824	1836384	3042014	2474774	5244878
24	1235674	537488	829381	229058	202690	787236	1137255	871717	2217329
25	675070	361386	466871	61949	47778	539070	613103	324351	1362024
26	189270	130877	163416	35156	29818	206295	259810	177234	424485
27	150294	109653	161369	31801	25917	180326	234863	167306	341679
28	108387	90111	138006	20276	24207	133383	219426	161044	306067
29	77843	81122	151905	13356	26787	109389	220000	152752	328587
30	39059	49499	107476	6368	24899	72613	174305	131087	231297
31	23740	33875	110670	3317	25948	56761	152283	117127	198552
32	16304	19851	97463	1883	26419	35390	166449	113240	157016

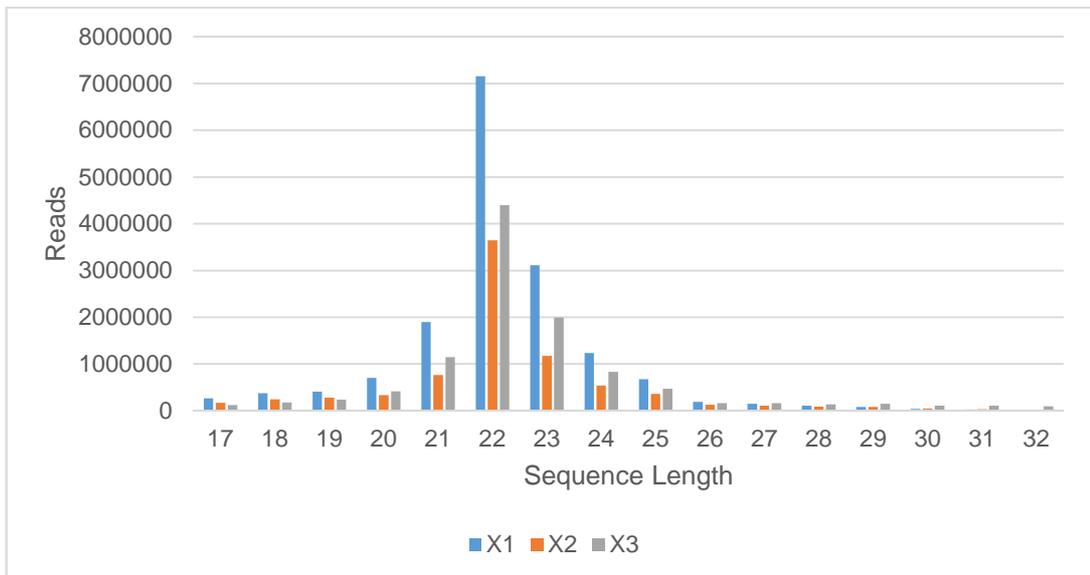


Figure 79: A graph representing the CCD-841 size class distribution data for X1, X2 and X3.

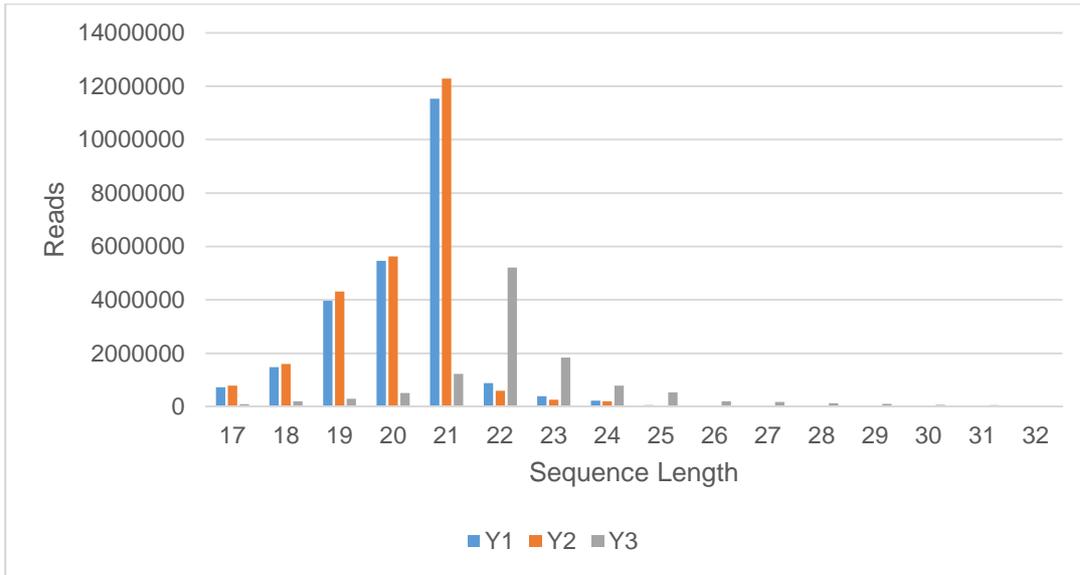


Figure 80: A graph representing the CCD-841 size class distribution data for Y1, Y2 and Y3.

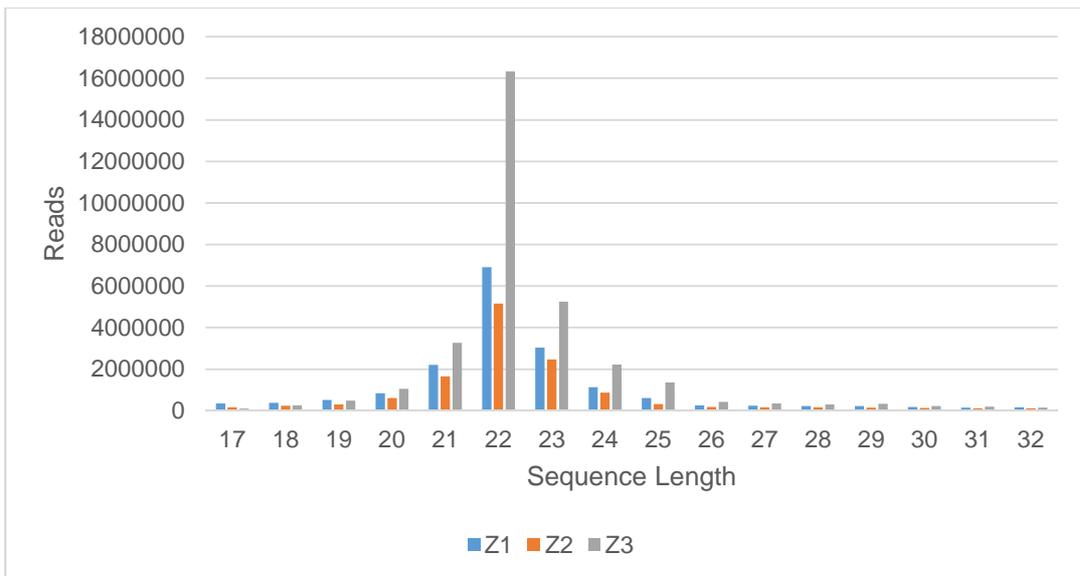


Figure 81: A graph representing the CCD-841 size class distribution data for Z1, Z2 and Z3.

The size distributions for CCD-841 libraries X1, X2, X3, Y3, Z1, Z2 and Z3 are as expected; Figure 79, Figure 80 and Figure 81 all show peaks for these at 22 nucleotides, as expected for human miRNAs. However, it is apparent from Figure 80 that Y1 and Y2 do not have the expected size class distributions, since they show peaks at 21 nucleotides, then sudden drops \geq 22 nucleotides, thus indicating issues with these libraries.

7.2.2.2 Complexity Checks

The size class distribution data presented here were generated by Claudia Paicu (School of Computing Sciences, University of East Anglia, Norwich, UK).

7.2.2.2.1 Caco-2 Libraries

Table 14: A summary of the complexity scores at each sequence length, in each Caco-2 library.

	A1	A2	A3	B1	B2	B3	C1	C2	C3
17	0.119	0.106	0.203	0.128	0.161	0.127	0.101	0.106	0.100
18	0.118	0.099	0.164	0.119	0.143	0.112	0.098	0.100	0.101
19	0.108	0.085	0.111	0.098	0.121	0.112	0.076	0.077	0.081
20	0.061	0.048	0.056	0.058	0.069	0.073	0.045	0.046	0.048
21	0.033	0.027	0.025	0.031	0.037	0.046	0.024	0.028	0.026
22	0.016	0.014	0.016	0.019	0.018	0.022	0.016	0.017	0.016
23	0.018	0.015	0.017	0.020	0.019	0.025	0.017	0.020	0.018
24	0.038	0.028	0.032	0.036	0.037	0.045	0.031	0.037	0.035
25	0.150	0.094	0.113	0.123	0.125	0.138	0.112	0.127	0.134
26	0.260	0.157	0.204	0.205	0.214	0.222	0.219	0.250	0.257
27	0.305	0.171	0.226	0.218	0.230	0.258	0.261	0.293	0.267
28	0.326	0.170	0.212	0.208	0.220	0.291	0.271	0.308	0.292
29	0.365	0.197	0.222	0.217	0.218	0.337	0.284	0.325	0.334
30	0.414	0.232	0.278	0.251	0.254	0.359	0.344	0.379	0.391
31	0.359	0.199	0.235	0.198	0.204	0.317	0.261	0.292	0.300
32	0.357	0.213	0.283	0.195	0.195	0.301	0.295	0.330	0.326

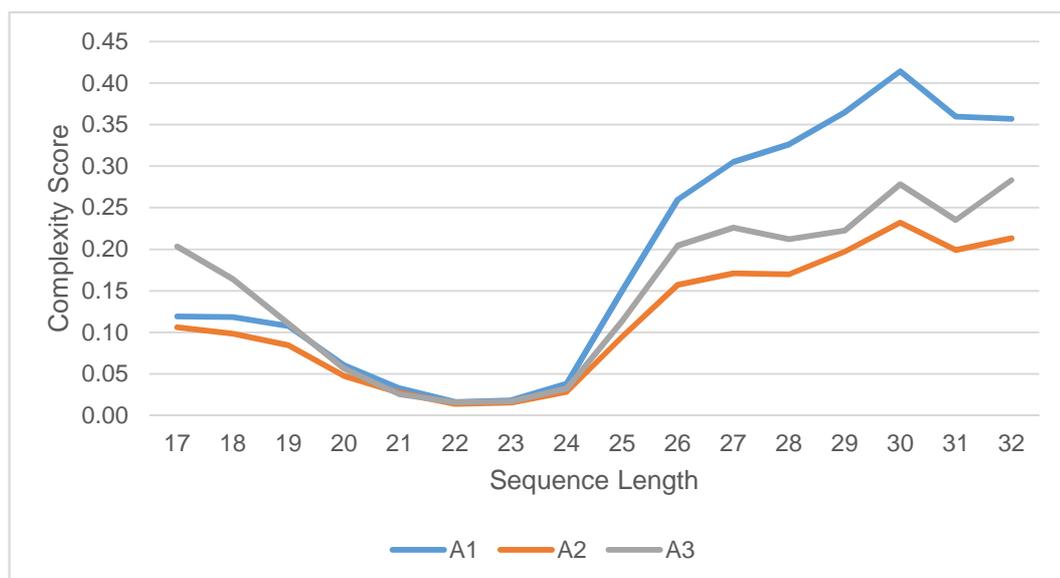


Figure 82: A graph illustrating the complexity distribution data for A1, A2 and A3.

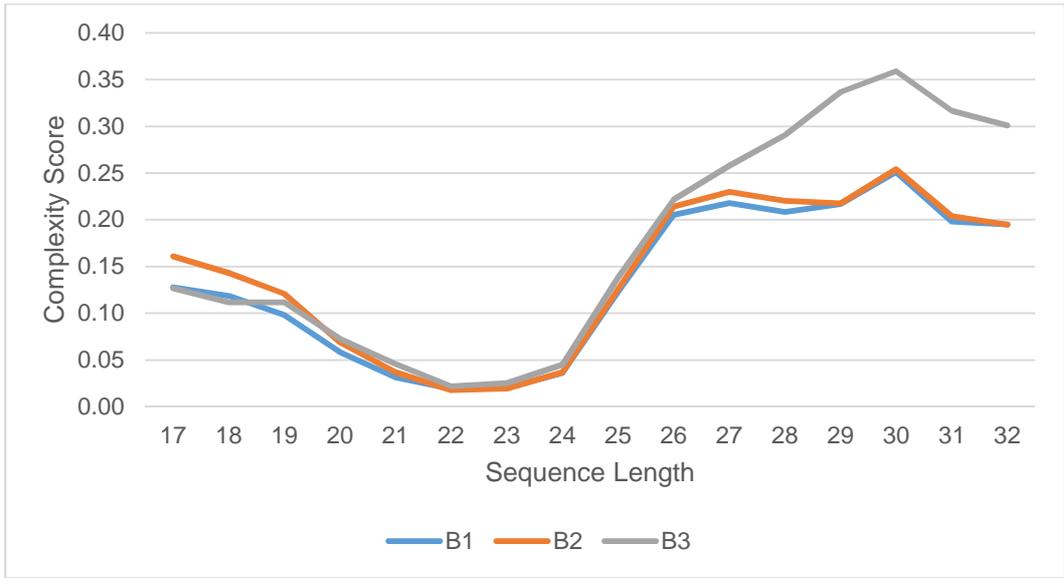


Figure 83: A graph illustrating the complexity distribution data for B1, B2 and B3.

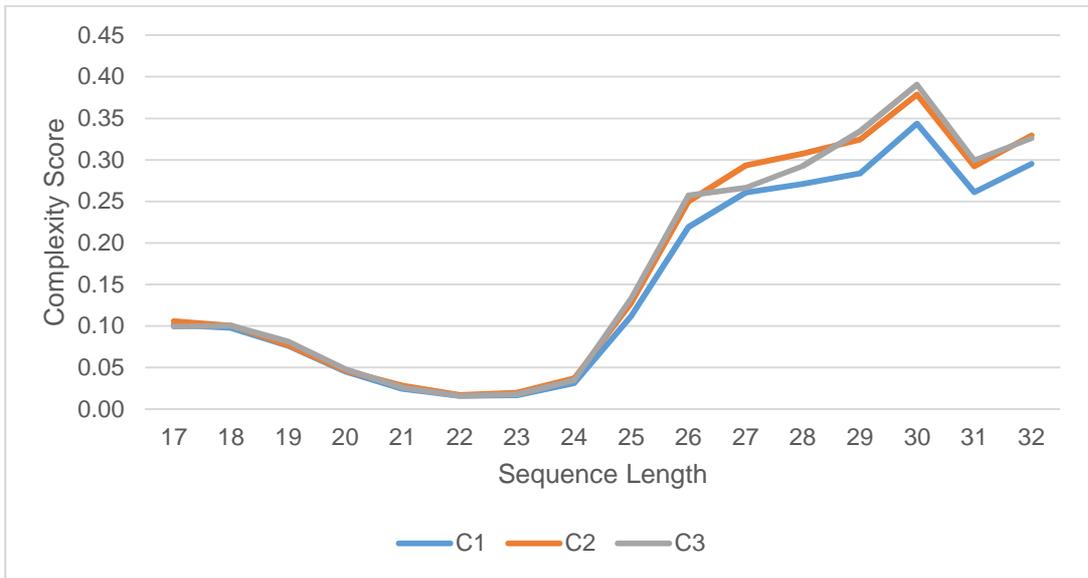


Figure 84: A graph illustrating the complexity distribution data for C1, C2 and C3.

The complexity distribution refers to the sequence complexity (i.e. how many unique sequences there are) at each given length of sequence. It is expected that there will inevitably be some random RNA degradation products present, distributed randomly per length. Human miRNAs tend to be between 18 and 25 nucleotides in length. The sequence variation of actual miRNAs will be

lower than that of random degradation products, so there should be a trough in complexity between 18 and 25 nucleotides, centred around 22. This expected trough is clearly visible for all of the Caco-2 libraries according to Figure 82, Figure 83 and Figure 84, suggesting that the data do not have too much noise from degradation products.

7.2.2.2.2 CCD-841 Libraries

Table 15: A summary of the complexity scores at each sequence length, in each CCD-841 library.

	X1	X2	X3	Y1	Y2	Y3	Z1	Z2	Z3
17	0.101	0.118	0.107	0.013	0.009	0.131	0.086	0.137	0.116
18	0.104	0.102	0.113	0.007	0.005	0.107	0.103	0.135	0.082
19	0.097	0.092	0.108	0.003	0.002	0.099	0.086	0.124	0.065
20	0.059	0.077	0.073	0.002	0.002	0.068	0.059	0.073	0.039
21	0.025	0.038	0.034	0.001	0.001	0.033	0.026	0.031	0.016
22	0.008	0.009	0.010	0.016	0.020	0.009	0.010	0.012	0.004
23	0.017	0.028	0.023	0.039	0.049	0.026	0.022	0.023	0.012
24	0.036	0.053	0.055	0.126	0.144	0.054	0.051	0.056	0.024
25	0.054	0.065	0.068	0.114	0.115	0.067	0.081	0.127	0.032
26	0.165	0.162	0.172	0.146	0.135	0.162	0.175	0.209	0.092
27	0.180	0.174	0.163	0.142	0.146	0.169	0.176	0.206	0.108
28	0.198	0.187	0.177	0.184	0.145	0.207	0.172	0.200	0.113
29	0.198	0.179	0.147	0.203	0.129	0.221	0.155	0.189	0.099
30	0.242	0.231	0.182	0.272	0.133	0.263	0.170	0.186	0.123
31	0.231	0.239	0.159	0.314	0.130	0.223	0.149	0.170	0.121
32	0.218	0.245	0.143	0.387	0.125	0.208	0.102	0.134	0.113

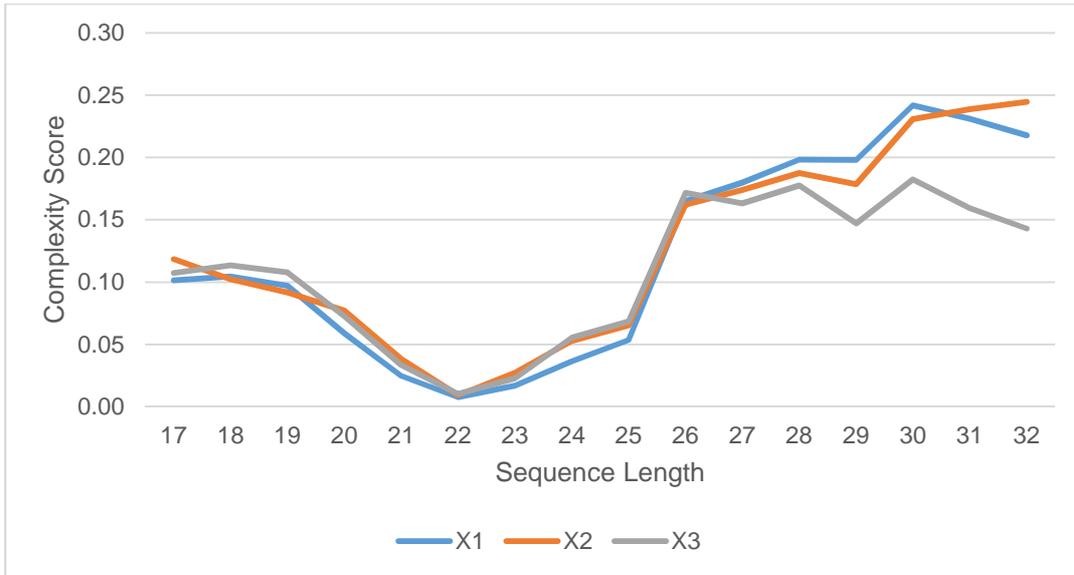


Figure 85: A graph illustrating the complexity distribution data for X1, X2 and X3.

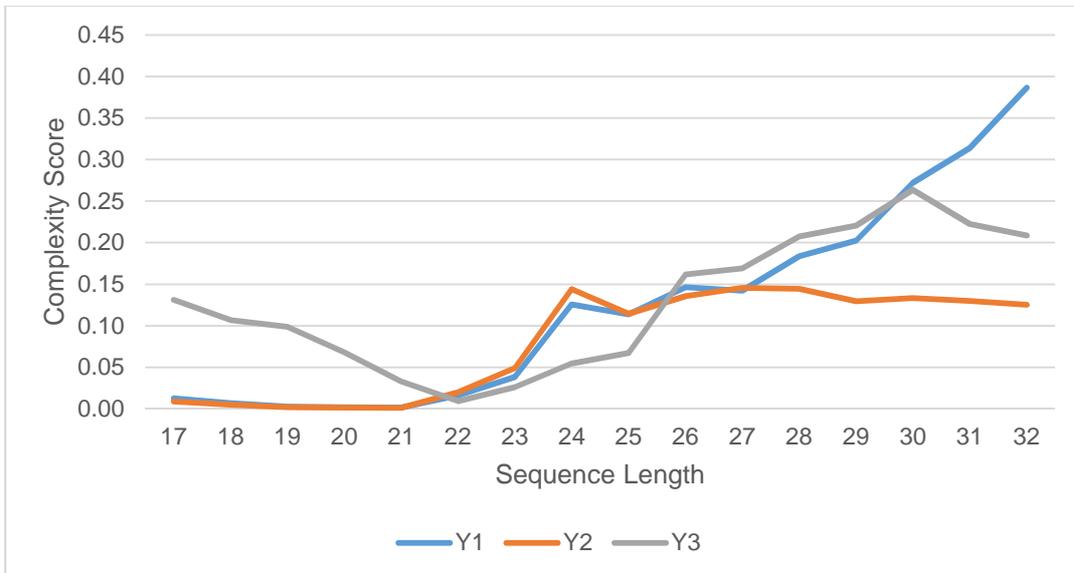


Figure 86: A graph illustrating the complexity distribution data for Y1, Y2 and Y3.

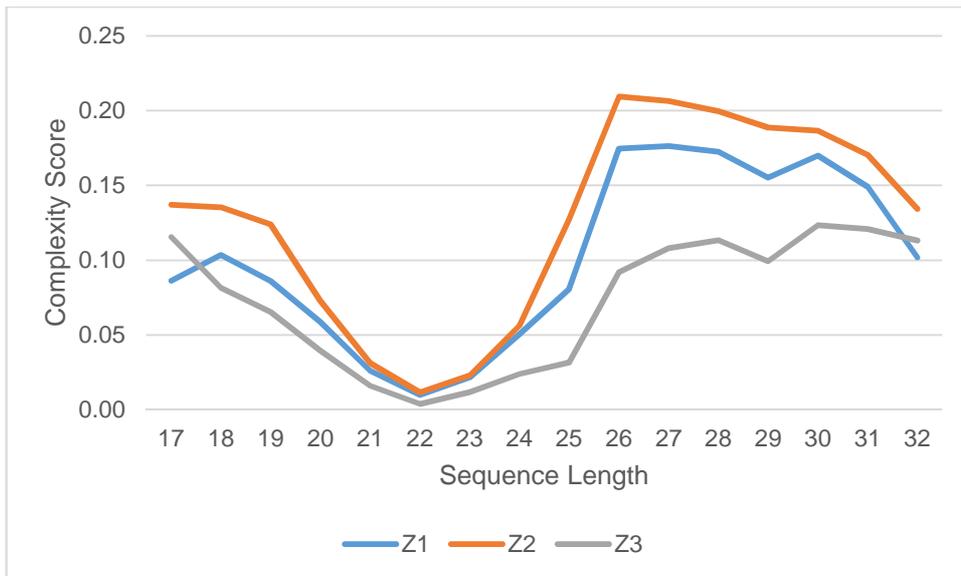


Figure 87: A graph illustrating the complexity distribution data for Z1, Z2 and Z3.

It is apparent from the CCD-841 library complexity distributions illustrated in Figure 85, Figure 86 and Figure 87 that libraries X1, X2, X3, Y3, Z1, Z2 and Z3 have the expected distributions i.e. with troughs centred around the average human miRNA length, 22 nucleotides, indicating that there is not much too noise from RNA degradation products. However, libraries Y1 and Y2 do not have the expected complexity distributions at all, as visible in Figure 86; complexity appears to be very low ≤ 21 nucleotides and then relatively high between 24 and 32 nucleotides. This suggests that these data may have a very high ratio of noise from RNA degradation products and are thus probably not useful.

7.2.3 Northern Blot/RT-qPCR Data

7.2.3.1 MicroRNAs Not Confirmed by Northern Blots/RT-qPCR

7.2.3.1.1 *hsa-miR-10b-5p* in Caco-2

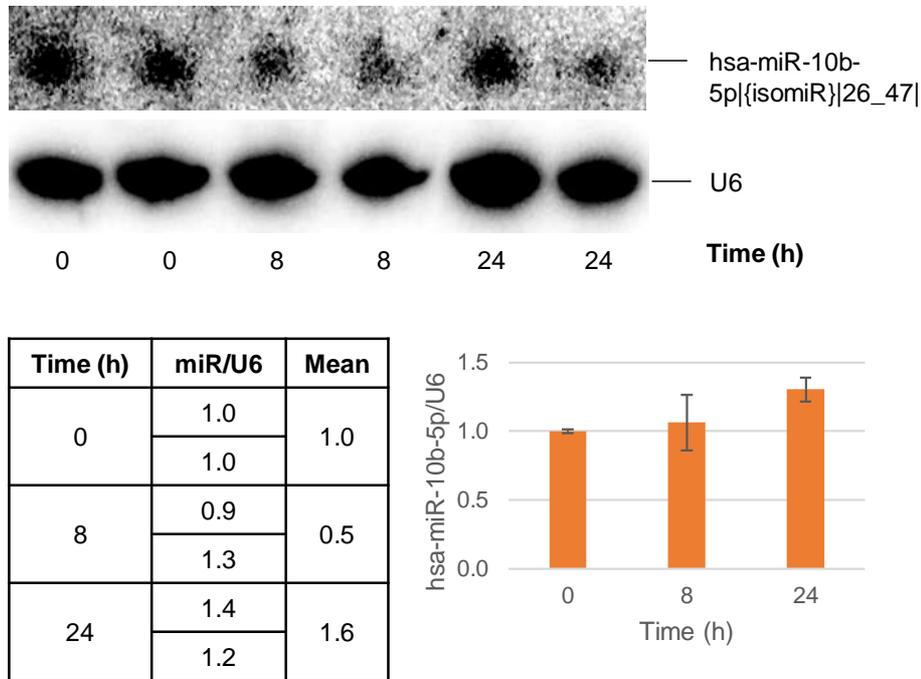


Figure 88: The effect of SFN treatment on the level of *hsa-miR-10b-5p* in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

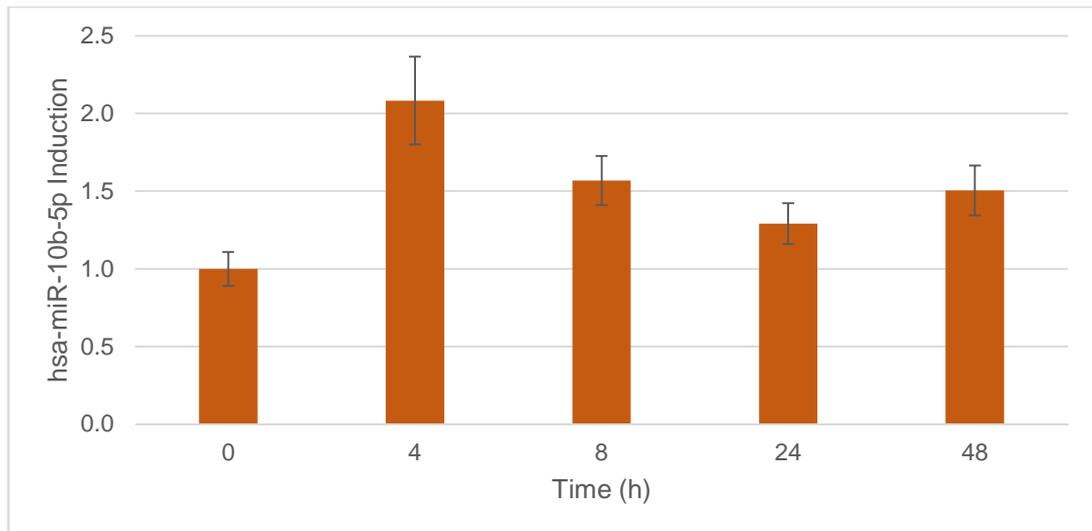


Figure 89: The effect of SFN treatment on the level of hsa-miR-10b-5p in Caco-2 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in triplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (*p<0.05).

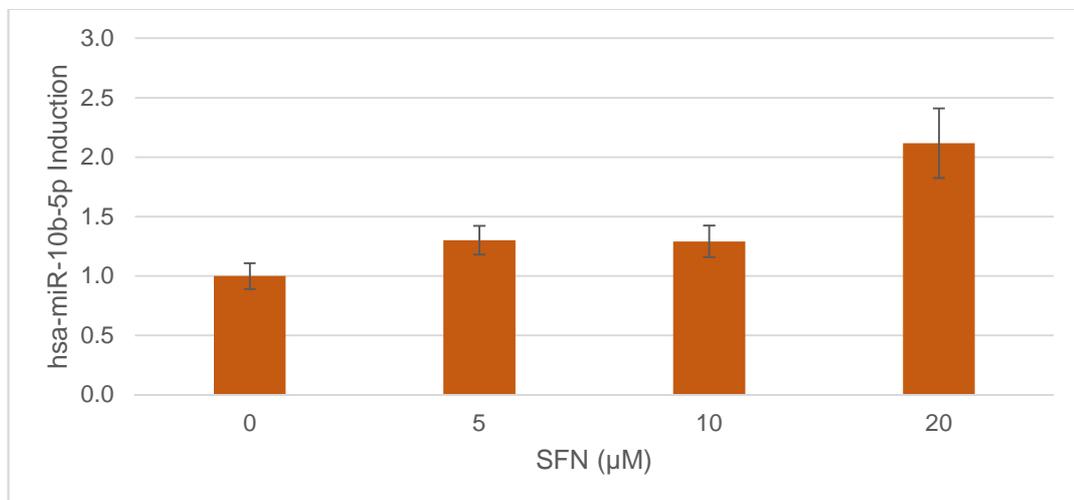


Figure 90: The effect of 24 h SFN treatment on the level of hsa-miR-10b-5p in Caco-2 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20 μ M) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (*p<0.05).

7.2.3.1.2 *hsa-miR-17-3p* in Caco-2

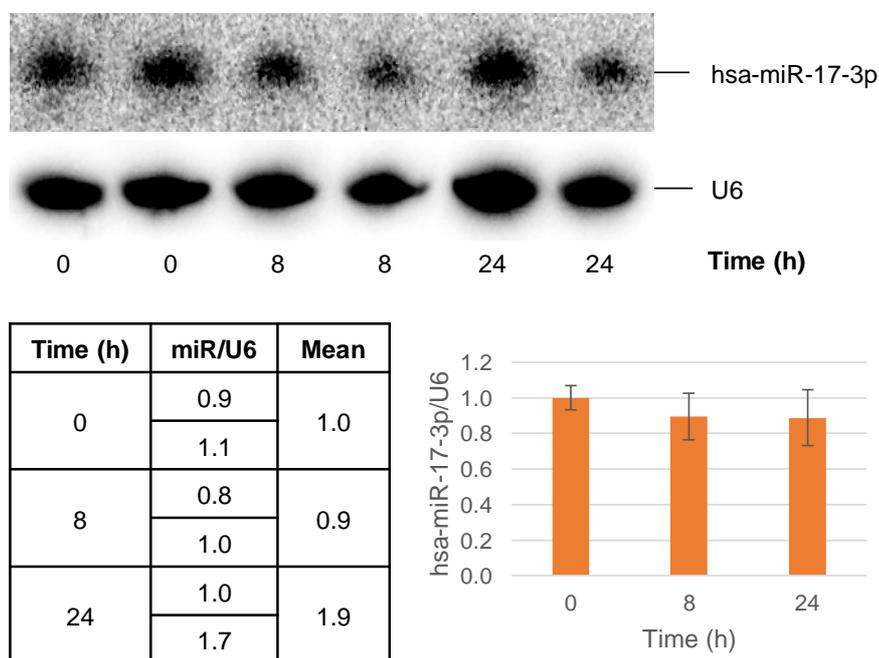


Figure 91: The effect of SFN treatment on the level of *hsa-miR-17-3p* in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.2.3.1.3 *hsa-miR-182-5p* in Caco-2

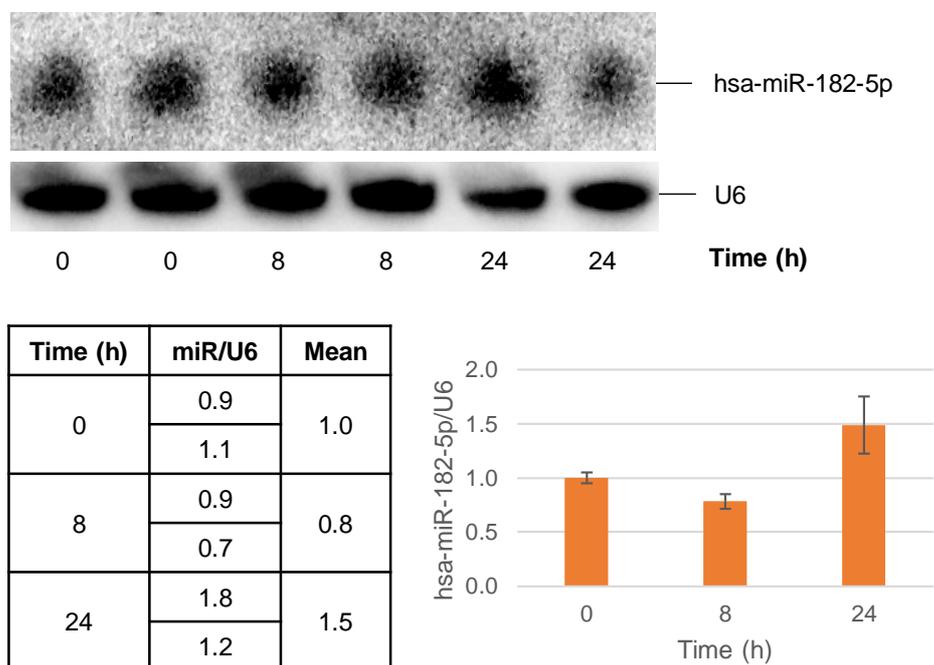


Figure 92: The effect of SFN treatment on the level of *hsa-miR-182-5p* in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.2.3.1.4 *hsa-miR-192-5p* in Caco-2

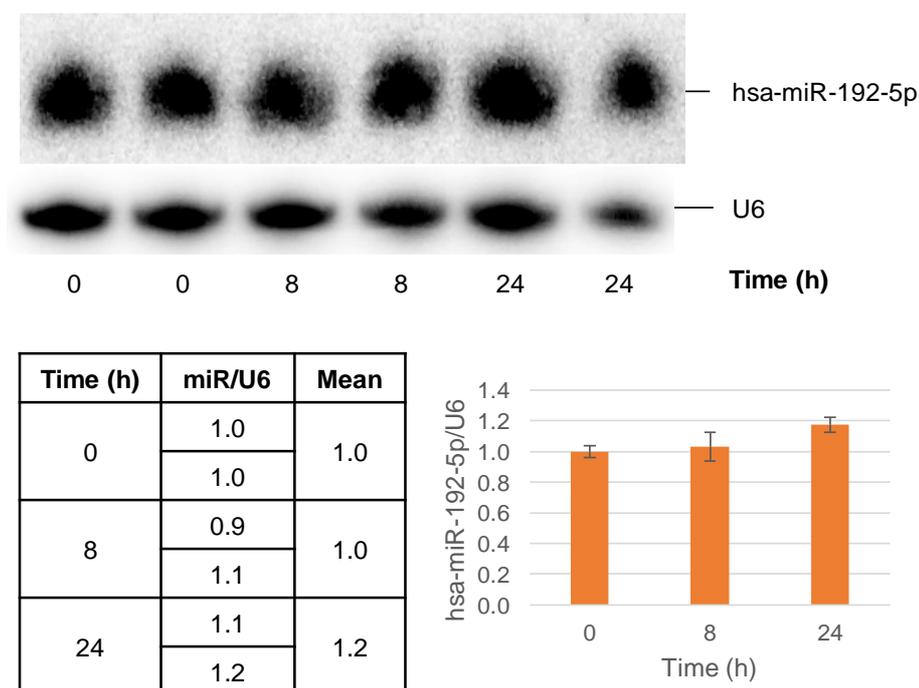


Figure 93: The effect of SFN treatment on the level of *hsa-miR-192-5p* in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.2.3.1.5 *hsa-miR-1296-5p* in Caco-2

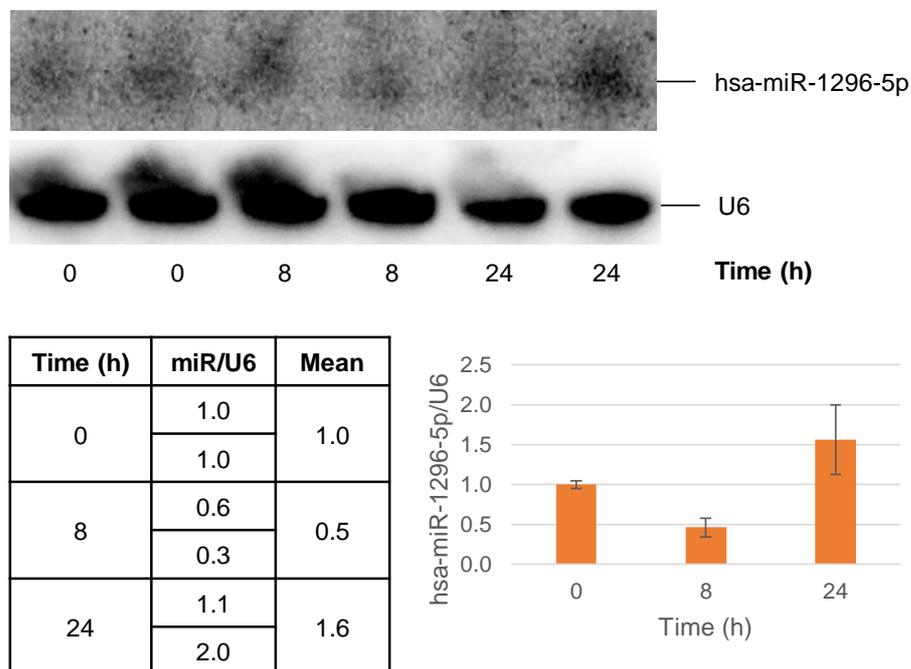
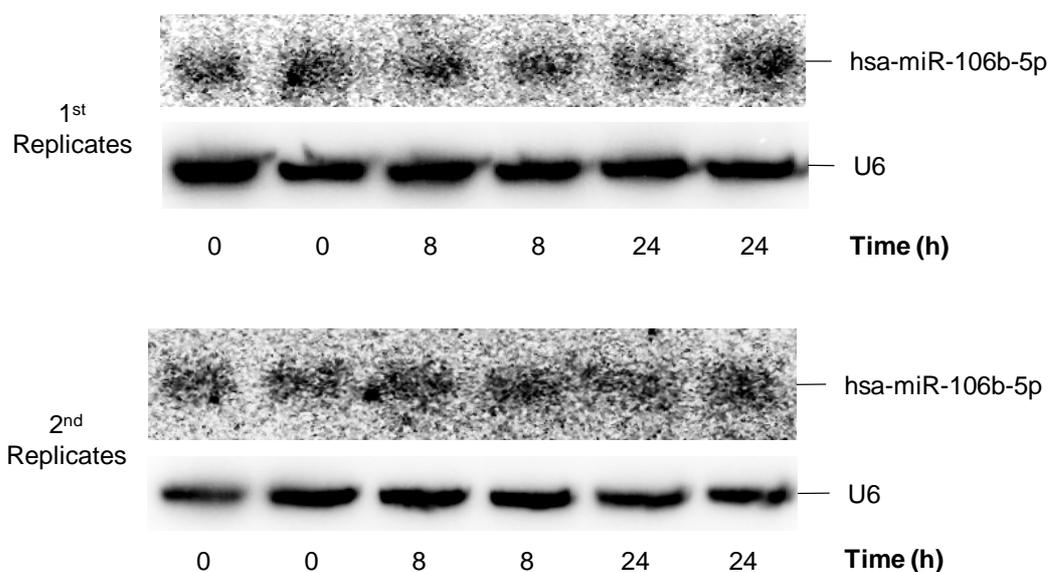


Figure 94: The effect of SFN treatment on the level of *hsa-miR-1296-5p* in Caco-2 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.2.3.1.6 *hsa-miR-106b-5p* in CCD-841



Time (h)	miR/U6 [1 st]	miR/U6 [2 nd]	Mean
0	1.0	1.0	1.0
4	1.9	0.7	1.3
8	1.3	1.1	1.2
12	1.2	0.7	1.0
24	1.3	0.8	1.0
48	2.1	1.2	1.6

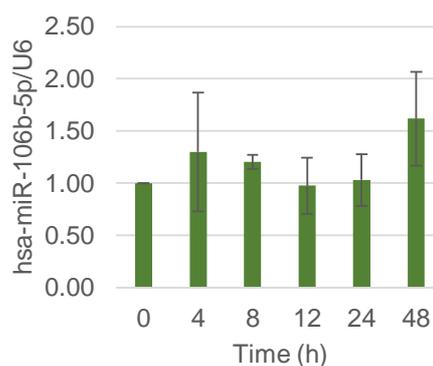


Figure 95: The effect of SFN treatment on the level of *hsa-miR-106b-5p* in CCD-841 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.2.3.1.7 *hsa-miR-345-5p* in CCD-841

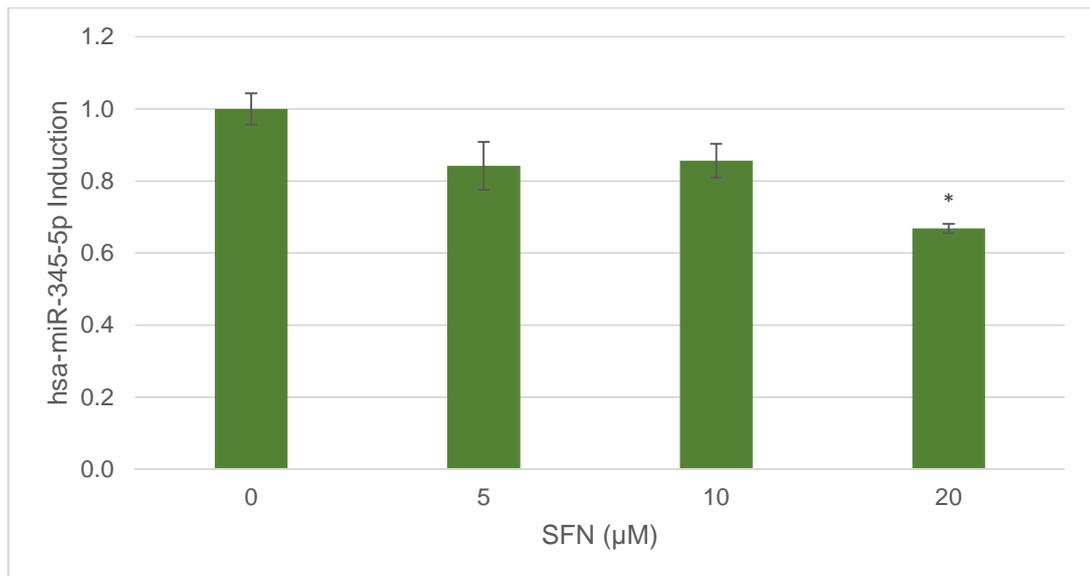
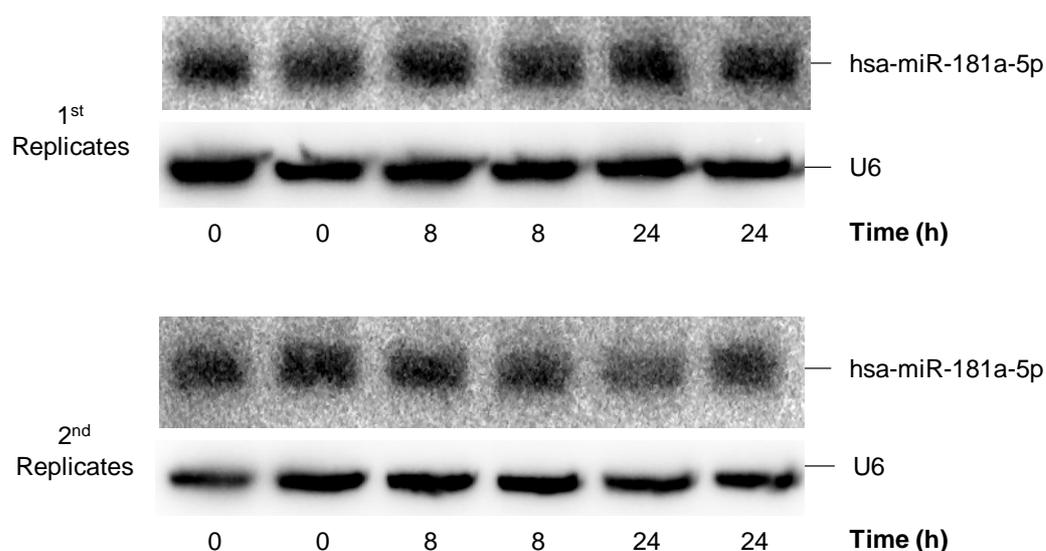


Figure 96: The effect of 24 h SFN treatment on the level of *hsa-miR-345-5p* in CCD-841 cells. Cells were treated for 24 h by adding to culture medium in triplicate, DMSO-diluted SFN (5, 10 or 20µM) or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via TaqMan RT-qPCR as described in Methods, using U6 as an internal control. Induction was evaluated by the Pfaffl method as described in Methods, and data are represented as means of triplicates \pm S.E.M. (*P<0.05).

7.2.3.1.8 *hsa-miR-181a-5p* in CCD-841



Time (h)	miR/U6 [1 st]	miR/U6 [2 nd]	Mean
0	1.0	1.0	1.0
4	1.3	0.8	1.0
8	1.1	0.8	0.9
12	1.1	0.8	0.9
24	1.0	0.8	0.9
48	1.0	1.1	1.0

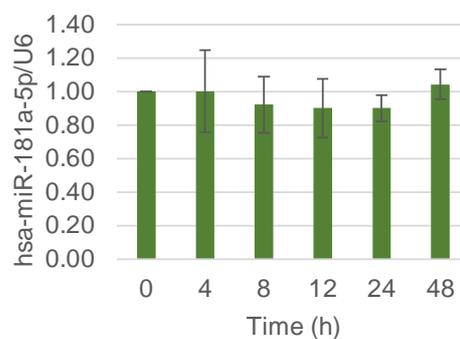


Figure 97: The effect of SFN treatment on the level of *hsa-miR-181a-5p* in CCD-841 cells at 4, 8, 12, 24 and 48 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 4, 8, 12, 24 or 48 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

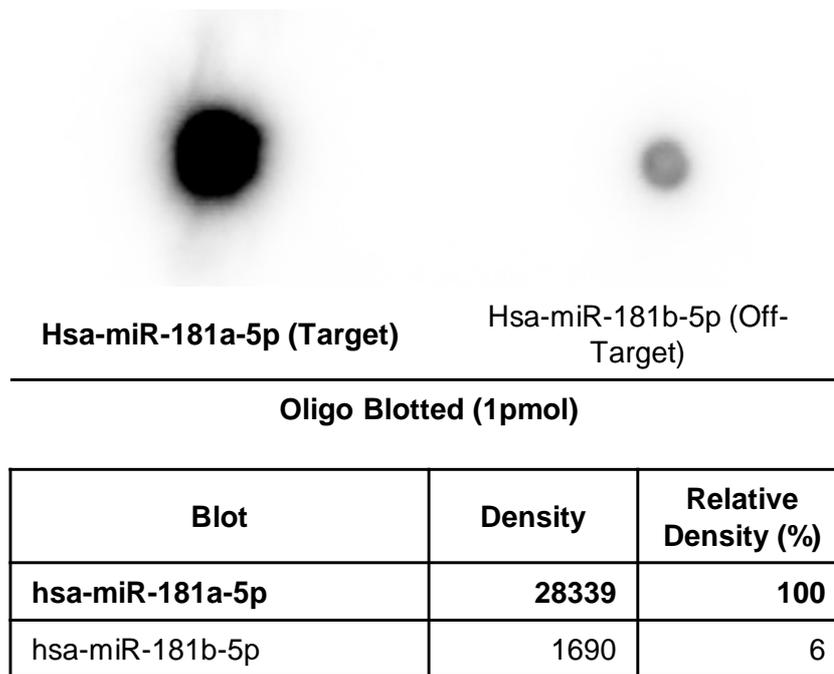


Figure 98: A test of the cross-reactivity of the probe anti-sense to hsa-miR-181a-5p with hsa-miR-181b-5p. Positive sense oligomers representing the sequence of each miRNA were blotted onto a membrane, which was then probed against hsa-miR-181a-5p

7.2.3.1.9 *hsa-miR-182-5p* in CCD-841

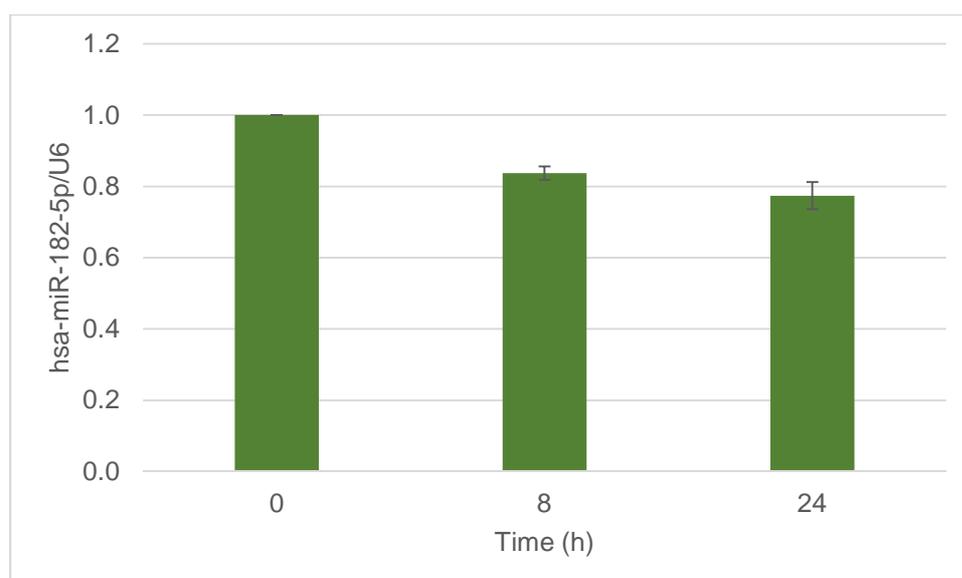


Figure 99: The effect of SFN treatment on the level of *hsa-miR-182-5p* in CCD-841 cells at 8 and 24 h. Cells were treated by adding to culture medium in duplicate, DMSO-diluted SFN (10 μ M) for 8 or 24 h, or DMSO (control); final DMSO concentrations were 0.05% (v/v). Relative miRNA abundance in RNA extracts was assayed via Northern Blotting as described in Methods, using U6 as an internal control. 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.

7.3 Additions to Chapter 2 and Chapter 5

7.3.1 Construction of pmiRGLO Vectors with 3' UTRs

7.3.1.1 Cloning the 3' UTRs

As described in Methods, parts of 3' UTRs containing predicted miRNA binding sites were cloned from human gDNA by PCR using designed primers. Listed below are the primers used for amplification, followed by the sequences of the amplified products (the predicted *hsa-let-7f-5p* binding sites are highlighted in yellow).

7.3.1.1.1 *BACH1*

Forward primer: CTAGTTGTTTAAACGGGTGTAGGGGGAGGATATTGC

Reverse primer: GACTCGAGGCTAGCGCACATGCAACCTGTATCGTCT

Sequence of the amplified product (1239bp):

GTGTAGGGGGAGGATATTGCTAGTATATTTTCAGTGGTTTGTATGTTCT
CTCTGTCACTGACTTATTTGTAAGAGAAAATTAGTTGGACTTGTTTATTTT
CTAGTAGCTTTTATAAGTACACTCAAGAATTTGTCAGGGAGAATAATTCT
GATAGTGCATCCCATACTGCAAAAGAATTTGTGTGTGTGTGTGTGTGTG
TGTGTGTGTGTGTGTATGTGTATGTATACATATATATCTCTCCATATAGG
TATTTCTTTGATACTTGTAAATTTTAAATTTTCAAGCTTCACGATATAAAATAAT
ATAAGAACTTCTGGTTTACAAAATGTAAAATCTTAAGCCAATGGAACCCT
TGATTTCCCTACCTCAGTGTACACCCAACTATTGGTTGTATCAGTTTGTGT
ATGTGCAAATGTCAAATAATCTTTTGCTTTAATTGCTACTGTACTTGCTTT
GAAAGATTACCTACTATTTTATGATAAAATGTAGTTGTCTCCAGAGCTTA
AATATAATTTGTAAAGCACTTGGTTTAAATTTCTCTCTACCTATAAACAGT
TTAGCATTAAAGGTTTCTATTAATGACACAGAATTATTGGCCAAGTGTA
TTTCTTAAAATTTAGCATTACTTTAAATAGCCAGCATGTA**ATACAAGTAAC**
TACACTACCTCATATCTACATGATTTTCAAGTTGTAATGCAGATGGACAG
ATAAAAAAGATTTTACGTTTGTCTTTTGGCCATAAGTGGGAAAGTTTTCT
GTATATTGCATAGCATTACACATTTATGCCTATTTTAAACATTAACCTTCTAA
AGAAGTTTTTTCTAAGAAAATGTTTCAAGGCAATATTTTTTTTGGAGGCTG
CCGAAGACAAATGACAGGATTATGAGTATACAGTGTATGCCTTTTCTT
CATGCAGAATTTTGAATGTTTTTCAAGTTTGTATATTGCATATTCACATGAT
CATTGTTCACTATTTTATGAACTGGCCTTCTCAATGTTTGTATGATTTTTTA
AAAGCTGTTATGTTGAATTCAGTAAAATAACATTACCTATTTTTTTTCTT
ATTCAAATTCTGGAACTATAGCAAATAATTCGTTAAATTGTCATATTCAAA
ACAAATGTGGATACAGTCTTGGTTCTCCATCTGTAATTTTTTTTAAACAGTT
TGCTATAGCTTACTGCTTAACTAATTTTAAATAAGGAAATAAGTATGTTAG
ATGCAGTAGACGATACAGGTTGCATGTG

7.3.1.1.2 *CDC25A[a]*

Forward primer: CTAGTTGTTTAAACGGTGACATTTGGAGAGGGGGC

Reverse primer: GACTCGAGGCTAGCGACAGGGACAGAAGAGGCGTA

Sequence of the amplified product (108bp):

GTGACATTTGGAGAGGGGGCCTGGGACTTCCATGCCTTAAACCTACCT
CCCACACTCCCAAGGTTGGAGCCCAGGGCATCTTGCTGGCTACGCCTC
TTCTGTCCCTGT

7.3.1.1.3 CDC25A[b]

Forward primer: CTAGTTGTTTAAACGCAGGGCATCTTGCTGGCTAC

Reverse primer: GACTCGAGGCTAGCGGGCAGAGAGCATGGGTCAA

Sequence of the amplified product (1613bp):

CAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCTGTTAGACGTCCTCC
GTCCATATCAGAAGTGTGCCACAATGCAGTTCTGAGCACCGTGTCAAGC
TGCTCTGAGCCACAGTGGGATGAACCAGCCGGGGCCTTATCGGGCTCC
AGCCATCTCATGAGGGGAGAGGAGACGGAGGGGAGTAGAGAAGTTAC
ACAGAAATGCTGCTGGCCAAATAGCAAAGACAACCTGGGAAGGAAAGG
TCTTTGTGGGATAATCCATATGTTTAAATTTATTCAACTTCATCAATCACTT
TATTTTATTTTTTTTTCTAACTCCTGGAGACTTATTTTACTGCTTCATTAG
GTTGAAATACTGCCATTCTAGGTAGGGTTTTATTATCCCAGGGACTACC
TCGGCTTTTAAATTTAAAAAAGAAAGTGGGTAAGAAAATGCAAACC
TGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCTGTCACCCCCAGGAG
GCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCCAAGAAGTACTGAGGTCC
TGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTGACATAAGCTAAATC
TCCCAGACCCACCACTGGCTACCGATATCTATTTGGTGGGAGGTGTGG
CCCTGTTCTTCCTCACCCAGTTCCATGACATTGGCTGGTATAGGAGCC
ACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTGGGCCACCCCGGCT
CAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGGGAAGGGGGTGGG
GGTAGGTGGGCTCCACAGGATGGGGGAGGAGCATGTCCACTGAGTAT
CTTCCTTATGTTGCTGTGATATTGATAGCTTTTATTTTCTAATTTTTAAAA
AATGGTCATATTATGAGTCAAAGAGTATCAAATCAGTGTTGGATGGACC
ACCAAGGGTGAGGAGAGGGGCTGGAAGCCCTGGGCATTAGGAGAAG
GGAGTGGGTGCTGGCATGGACATGACTGGATAGAATTTTCTCAGGAGG
GAGCTTGGTGGATTTTGAAGGTAAAATTTCTGGGTTTATCATGTTTTAA
TTTTAGAGACAGGGAGTGATGAATCATCACCGGTTGTCCCTTATCTAA
CTCCATAAAAGTGGGAATTTCAAAGAACACCTCATCCAAGGAGCTGGG
GCAGACTTCATTGATTCTAGAGAGACCTGTTTCAGTGCCTACTCATCCC

TGCCCTCTGGTGCCAGCCTCCTTACCATCACGGCTTCACTGAGGTGTA
GGTGGGTTTTCTTAAACAGGAGACAGTCTCTCCCCTCTTACCTCAACT
TCTTGGGGTGGGAATCAGTGATACTGGAGATGGCTAGTTGCTGTGTTAC
GGGTTTGAGTTACATTTGGCTATAAAACAATCTTGTTGGGAAAATGTG
GGGGAGAGGACTTCTTCCTACACGCGCATTGAGACAGATTCCAAGTGG
TTAATGATATTGTTTGTAAGAAAGAGATTCTGTTGGTTGACTGCCTAAG
AGAAAGGTGGGATGGCCTTCAGATTATACCAGCTTAGCTAGCATTACTA
ACCAACTGTTGGAAGCTCTGAAAATAAAAGATCTTGAACCCATGCTCTC
TGCC

7.3.1.1.4 *HMGA2[a] (Failed, Re-Done)*

Forward primer: CTAGTTGTTTAAACGGTTCGATTTCTACCTCAGCAGC

Reverse primer: GACTCGAGGCTAGCGCGTGTTCCCTTCTATCAAATGTC

Sequence of the amplified product (1793bp):

GTTTCGATTTCTACCTCAGCAGCAGTTGGATCTTTTGAAGGGAGAAGACA
CTGCAGTGACCACTTATTCTGTATTGCCATGGTCTTTCCACTTTCATCTG
GGGTGGGGTGGGGTGGGGTGGGGGAGGGGGGGGGTGGGGTGGGGAG
AAATCACATAACCTTAAAAAGGACTATATTAATCACCTTCTTTGTAATCCC
TTCACAGTCCCAGGTTTAGTGAAAACTGCTGTAAACACAGGGGACACA
GCTTAACAATGCAACTTTTAATTACTGTTTTCTTTTTCTTAACCTACTAA
TAGTTTGTTGATCTGATAAGCAAGAGTGGGCGGGTGAGAAAAACCGAAT
TGGGTTTAGTCAATCACTGCACTGCATGCAAACAAGAAACGTGTCACAC
TTGTGACGTCGGGCATTCATATAGGAAGAACGCGGTGTGTAACACTGT
GTACACCTCAAATACCACCCCAACCCACTCCCTGTAGTGAATCCTCTGT
TTAGAACACCAAAGATAAGGACTAGATACTACTTTCTTTTTTCGTATAA
TCTTGTAGACACTTACTTGATGATTTTTAACTTTTTATTTCTAAATGAGAC
GAAATGCTGATGTATCCTTTCATTCAGCTAACAACTAGAAAAGGTTATG
TTCATTTTTCAAAAAGGGAAGTAAGCAAACAATATTGCCAACTCTTCTA
TTTATGGATATCACACATATCAGCAGGAGTAATAAATTTACTCACAGCAC
TTGTTTTTCAGGACAACACTTCATTTTCAGGAAATCTACTTCCTACAGAGC
CAAATGCCATTTAGCAATAAATAACACTTGTCAGCCTCAGAGCATTTAA
GGAACTAGACAAGTAAAATTATCCTCTTTGTAATTTAATGAAAAGGTAC
AACAGAATAATGCATGATGAACTCACCTAATTATGAGGTGGGAGGAGCG

AAATCTAAATTTCTTTTGCTATAGTTATACATCAATTTAAAAAGCAAAAA
AAAAAGGGGGGGCAATCTCTCTGTGTCTTTCTCTCTCTCTCTCC
TCTCCCTCTCTTTTCATTGTGTATCAGTTTCCATGAAAGACCTGAATA
CCTACTTACCTCAAATTAAGCATATGTGTTACTTCAAGTAATACGTTTTGA
CATAAGATGGTTGACCAAGGTGCTTTTCTTCGGCTTGAGTTCACCATCT
CTTCATTCAAACCTGCACTTTTAGCCAGAGATGCAATATATCCCCACTACT
CAATACTACCTCTGAATGTTACAACGAATTTACAGTCTAGTACTTATTAC
ATGCTGCTATACACAAGCAATGCAAGAAAAAACTTACTGGGTAGGTGA
TTCTAATCATCTGCAGTTCTTTTTGTACACTTAATTACAGTTAAAGAAGCA
ATCTCCTTACTGTGTTTCAGCATGACTATGTATTTTTCTATGTTTTTTAA
TAAAAATTTTTAAATACTTGTTTCAGCTTCTCTGCTAGATTTCTACATT
AACTTGAAAATTTTTAAACCAAGTCGCTCCTAGGTTCTTAAGGATAATTTT
CCTCAATCACACTACACATCACACAAGATTTGACTGTAATATTTAAATATT
ACCCCTCCAAGTCTGTACCTCAAATGAATCTTTAAGGAGATGGACTAATT
GACTTGCAAAGACCTACCTCAGACTTCAAAGGAATGAACTTGTTACT
TGCAGCATTCAATTTGTTTTTCAATGTTTGAAATAGTTCAAACCTGCAGCT
AACCTAGTCAAACCTATTTTTGTAAAAGACATTTGATAGAAAGGAACAC
G

7.3.1.1.5 *HMGA2[a] (Re-try with New Primers)*

Forward primer: CTAGTTGTTTAAACGGGGGCGCCAACGTTTCGATTT

Reverse primer: GACTCGAGGCTAGCGTGACTAGGGTTAGCTGCAGTTTG

Sequence of the amplified product (1764bp):

GGGGCGCCAACGTTTCGATTTCTACCTCAGCAGCAGTTGGATCTTTTGAA
GGGAGAAGACACTGCAGTGACCACTTATTCTGTATTGCCATGGTCTTTT
CACTTTCATCTGGGGTGGGGTGGGGTGGGGTGGGGGAGGGGGGGGT
GGGGTGGGGAGAAATCACATAACCTTAAAAAGGACTATATTAATCACCT
TCTTTGTAATCCCTTCACAGTCCCAGGTTTAGTGAAAACTGCTGTAAAC
ACAGGGGACACAGCTTAACAATGCAACTTTTAATTAATGTTTTCTTTTT
CTTAACCTACTAATAGTTTGTTGATCTGATAAGCAAGAGTGGGCGGGTG
AGAAAACCGAATTGGGTTTAGTCAATCACTGCACTGCATGCAAACAAG
AAACGTGTCACACTTGTGACGTCGGGCATTCATATAGGAAGAACGCGG
TGTGTAACACTGTGTACACCTCAAATACCACCCAACCCACTCCCTGTA

GTGAATCCTCTGTTTAGAACACCAAAGATAAGGACTAGATACTACTTTCT
CTTTTTCGTATAATCTTGTAGACACTTACTTGATGATTTTTAACTTTTTATT
TCTAAATGAGACGAAATGCTGATGTATCCTTTCATTCAGCTAACAACTA
GAAAAGGTTATGTTCATTTTTCAAAAAGGGAAGTAAGCAAACAATATTG
CCAACCTCTTCTATTTATGGATATCACACATATCAGCAGGAGTAATAAATT
TACTCACAGCACTTGTTTTCAGGACAACACTTCATTTTTCAGGAAATCTAC
TTCCTACAGAGCCAAAATGCCATTTAGCAATAAATAACACTTGTCAGCCT
CAGAGCATTTAAGGAACTAGACAAGTAAAATTATCCTCTTTGTAATTTA
ATGAAAAGGTACAACAGAATAATGCATGATGAACTCACCTAATTATGAG
GTGGGAGGAGCGAAATCTAAATTTCTTTTGCTATAGTTATACATCAATTT
AAAAAGCAAAAAAAAAAAGGGGGGGCAATCTCTCTCTGTGTCTTTCT
CTCTCTCTTCTCTCCCTCTCTCTTTTCATTGTGTATCAGTTTCCATGA
AAGACCTGAATACCACCTTACCTCAAATTAAGCATATGTGTTACTTCAAGT
AATACGTTTTGACATAAGATGGTTGACCAAGGTGCTTTTCTTCGGCTTGA
GTTACCATCTCTTCATTCAACTGCACTTTTAGCCAGAGATGCAATATA
TCCCCACTACTCAATACTACCTCTGAATGTTACAACGAATTTACAGTCTA
GTACTTATTACATGCTGCTATACACAAGCAATGCAAGAAAAAACTTACT
GGTAGGTGATTCTAATCATCTGCAGTTCTTTTTGTACTTAATTACAG
TTAAGAAGCAATCTCCTTACTGTGTTTCAGCATGACTATGTATTTTTCTA
TGTTTTTTAATTAATAAATTTTTAAATACTTGTTTCAGCTTCTCTGCTAGA
TTTCTACATTACTTGAAAATTTTTAAACCAAGTCGCTCCTAGGTTCTTAA
GGATAATTTTCCTCAATCACACTACACATCACACAAGATTTGACTGTAAT
ATTTAAATATTACCCCTCCAAGTCTGTACCTCAAATGAATCTTTAAGGAG
ATGGACTAATTGACTTGCAAAGACCTACCTCAGACTTCAAAGGAATG
AACTTGTTACTTGCAGCATTCAATTTGTTTTTCAATGTTTGAATAGTTCA
AACTGCAGCTAACCCCTAGTCA

7.3.1.1.6 *HMGA2*[b]

Forward primer: CTAGTTGTTTAAACGGCTGCTTCAGGGAGGTAGTT

Reverse primer: GACTCGAGGCTAGCGGCGACCAACAACAGCAAAGAA

Sequence of the amplified product (479bp):

GCTGCTTCAGGGAGGTAGTTTCAAAGGCCACATACCTCTCTGAGACTG
GCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTATGGAGAGAATTA

AAACTCAACATTACTGTAACTGTGCGTTAAATAAGCAAATAAACAGTGG
CTCATAAAAATAAAAAGTCGCATTCCATATCTTTGGATGGGCCTTTTAGAA
ACCTCATTGGCCAGCTCATAAAATGGAAGCAATTGCTCATGTTGGCCAA
ACATGGTGCACCGAGTGATTTCCATCTCTGGTAAAGTTACACTTTTATTT
CCTGTATGTTGTACAATCAAAACACACTACTACCTCTTAAGTCCCAGTAT
ACCTCATTTTTTCATACTGAAAAAAAAAGCTTGTGGCCAATGGAACAGTAA
GAACATCATAAAATTTTTATATATATAGTTTATTTTTGTGGGAGATAAATT
TTATAGGACTGTTCTTTGCTGTTGTTGGTCGC

7.3.1.1.7 KRAS

Forward primer: CTAGTTGTTTAAACGGGTCTCTGTGCCAGCTCTA

Reverse primer: GACTCGAGGCTAGCGTGCCTACTTGGGAACATTCACT

Sequence of the amplified product (757bp):

GGTCTCTGTGCCAGCTCTATAATTGTTTTGCTACGATTCCACTGAAACTC
TTCGATCAAGCTACTTTATGTAAATCACTTCATTGTTTTAAAGGAATAAAC
TTGATTATATTGTTTTTTTATTTGGCATAACTGTGATTCTTTTAGGACAAT
TACTGTACACATTAAGGTGTATGTCAGATATTCATATTGACCCAAATGTG
TAATATTCCAGTTTTCTCTGCATAAGTAATTTAAATATACTTAAAAATTAA
TAGTTTTATCTGGGTACAAATAAACAGGTGCCTGAACTAGTTCACAGAC
AAGGAACTTCTATGTAAAAATCACTATGATTTCTGAATTGCTATGTGAA
ACTACAGATCTTTGGAACACTGTTTAGGTAGGGTGTTAAGACTTACACA
GTACCTCGTTTCTACACAGAGAAAGAAATGGCCATACTTCAGGAACTGC
AGTGCTTATGAGGGGATATTTAGGCCTCTTGAATTTTTGATGTAGATGG
GCATTTTTTTAAGGTAGTGGTTAATTACCTTTATGTGAACTTTGAATGGTT
TAACAAAAGATTTGTTTTGTAGAGATTTTAAAGGGGAGAATTCTAGAA
ATAAATGTTACCTAATTATTACAGCCTTAAAGACAAAAATCCTTGTTGAA
GTTTTTTTTAAAAAAGCTAAATTACATAGACTTAGGCATTAACATGTTTGT
GGAAGAATATAGCAGACGTATATTGTATCATTTGAGTGAATGTTCCCAA
GTAGGCA

7.3.1.1.8 MYC

Forward primer: CTAGTTGTTTAAACGAATGTCCTGAGCAATCAC

Reverse primer: GACTCGAGGCTAGCGGAGGCAGTTTACATTATGGC

Sequence of the amplified product (112bp):

AATGTCCTGAGCAATCACCTATGAACTTGTTTCAAATGCATGATCAAATG
 CAACCTCACAACTTGGCTGAGTCTTGAGACTGAAAGATTTAG **CCATAA**
TGTAAACTGCCTC

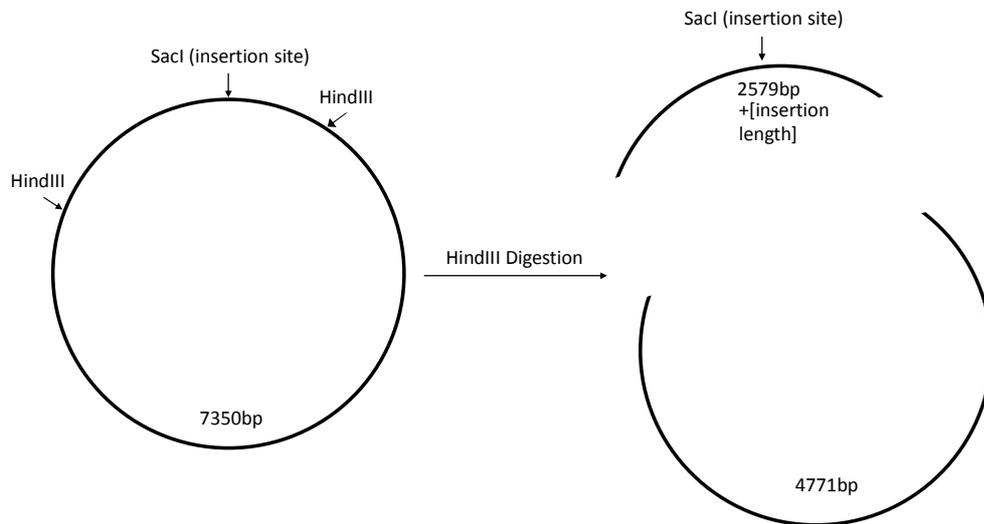


Figure 100: A diagram illustrating the expected cleavage of the pmiRGLO vector by the HindIII restriction enzyme.

7.3.1.2 Checking Prepared Plasmids for the Presence of Desired Inserts

Table 16: A summary of the expected HindIII digestion fragments.

Insert	Insert Length (bp)	Length of Fragment with Insertion Site (bp)	Length of Other Fragment (bp)
None	0	2579	4771
BACH1	1239	3818	4771
CDC25A[b]	1613	4192	4771
KRAS	757	3336	4771

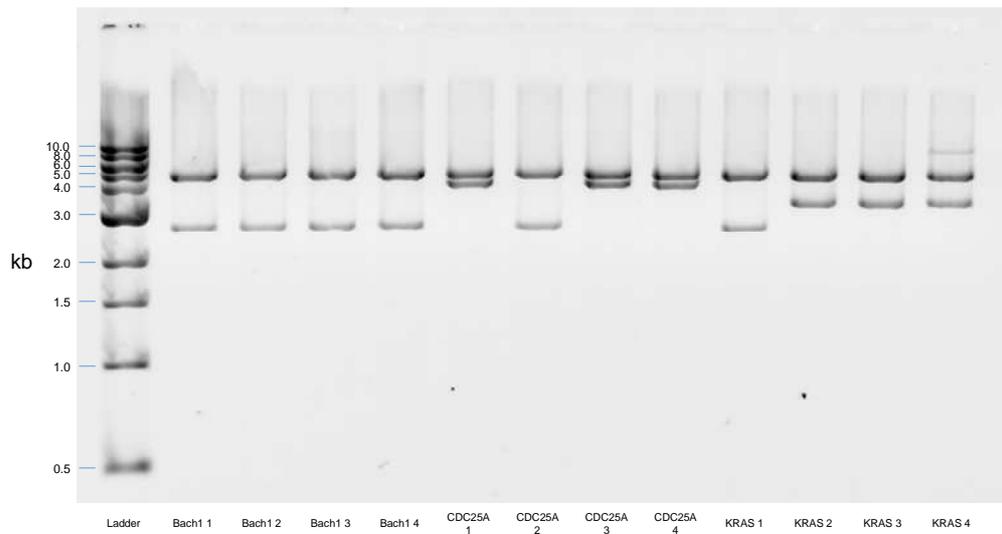


Figure 101: An ethidium bromide-based image of a 1% agarose gel into which HindIII digestion products for BACH1, BACH2, BACH3, BACH4, CDC25A[b] 1, CDC25A[b] 2, CDC25A[b] 3, CDC25A[b] 4, KRAS 1, KRAS 2, KRAS 3 and KRAS 4 were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2 h.

The band positions in Figure 101 clearly show that CDC25A 1, CDC25A 3, CDC25A 4, KRAS 2, KRAS 3 and KRAS 4 plasmids all have the desired inserts, while all of the other plasmids have no insert.

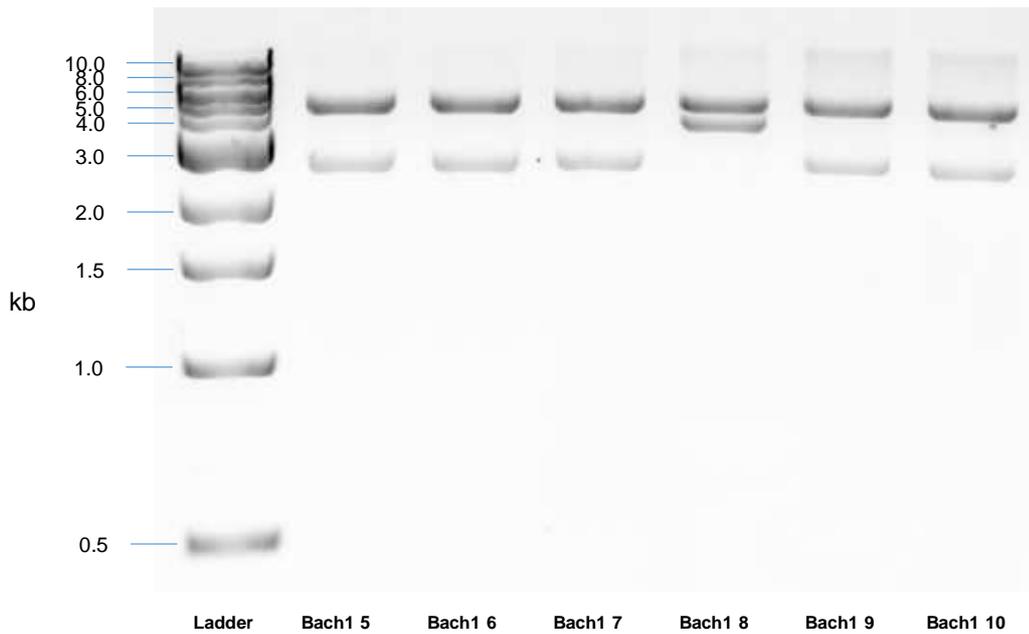


Figure 102: An ethidium bromide-based image of a 1% agarose gel into which HindIII digestion products for BACH5, BACH6, BACH7, BACH8, BACH9 and BACH10 were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2h.

According to the band positions in Figure 102, only BACH1 8 had the BACH1 3'-UTR insert.

Table 17: A summary of the expected HindIII digestion products. *The HMGA2[b] region has one HindIII digestion site within it.

Insert	Insert Length (bp)	Length of Fragment(s) with Insertion Site (bp)	Length of Other Fragment (bp)
None	0	2579	4771
HMGA2[a]	1793	4372	4771
HMGA2[b]	479	2080 and 978*	4771
MYC	112	2691	4771

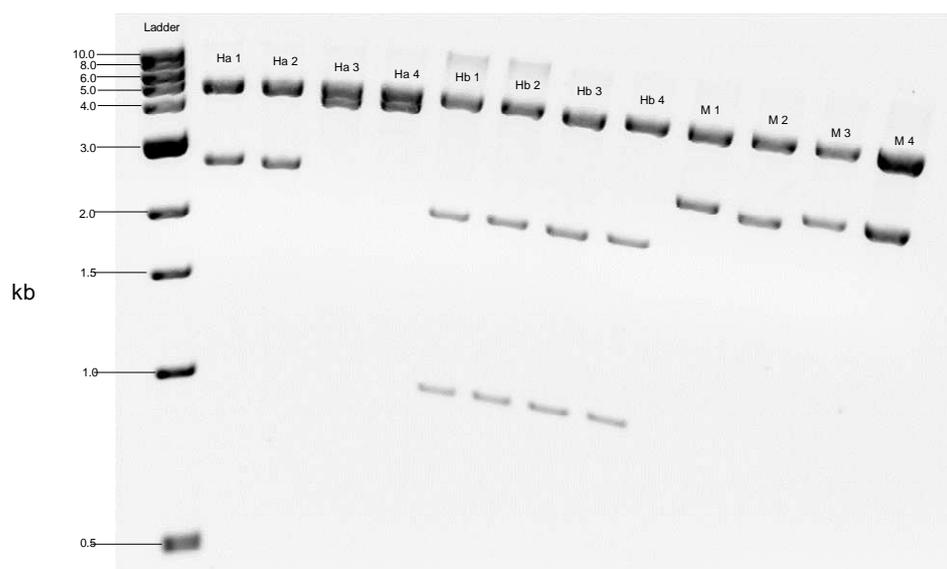


Figure 103: An ethidium bromide-based image of a 1% agarose gel into which the HindIII digestion products of the HMGA2[a] 1, HMGA2[a] 2, HMGA2[a] 3, HMGA2[a] 4, HMGA2[b] 1, HMGA2[b] 2, HMGA2[b] 3, HMGA2[b] 4, MYC 1, MYC 2, MYC 3 and MYC 4 plasmids were loaded along a 1kb ladder (NEB), and electrophoresed at 100V for 2 h 30 min.

Although the gel image shown in Figure 103 shows distortion, it is still clear that HMGA2[a] 3, HMGA[a] 4, and all four HMGA2[b] plasmids contain the desired inserts (however, it later became apparent that these HMGA2[a] constructs were unsuccessful). For the MYC inserts, it is unclear whether or not the insert is present due to the relatively short length of the insert.

Table 18: A summary of expected products following digestion with both HindIII and NheI.

Insert	Insert Length (bp)	Lengths of Fragments Following HindIII Digestion (bp)		Lengths of Fragments Following Subsequent NheI Digestion (bp)		
		2579	4771	867	<u>1710</u>	4771
None	0	2579	4771	867	<u>1710</u>	4771
CDC25A[a]	1793	4372	4771	867	<u>1818</u>	4771
MYC	112	2691	4771	867	<u>1822</u>	4771

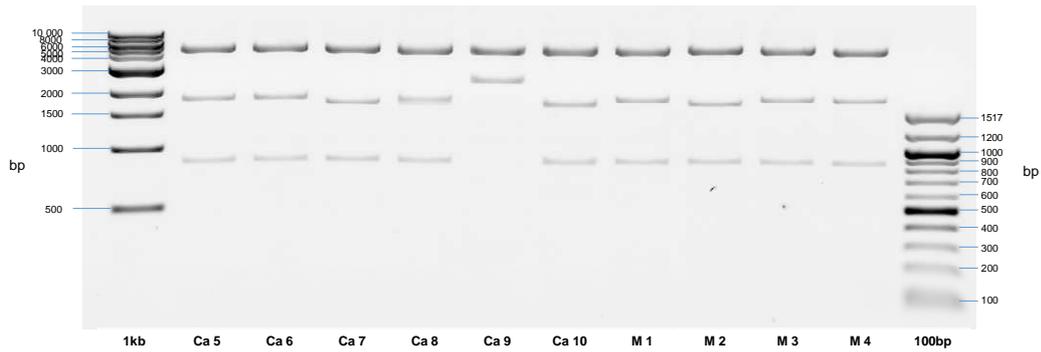


Figure 104: An ethidium bromide-based image of a 1% agarose gel into which the HindIII and NheI digestion products of the CDC25A[a] 5, CDC25A[a] 6, CDC25A[a] 7, CDC25A[a] 8, CDC25A[a] 9, MYC 1, MYC 2, MYC 3 and MYC 4 plasmids were loaded alongside 1kb and 100bp ladders (NEB), and electrophoresed at 80V for 1 h.

It is clear from the image in Figure 104 that the CDC25A[a] 5, CDC25A[a] 6 and CDC25A[a] 8 plasmids have the insert, whereas CDC25A[a] 7 does not. It is also clear that all of the MYC plasmids except for MYC 2 have the desired insert.

The artificial “let-7f-5p antisense repeat” insert was 118bp in length, so if present in the plasmid, HindIII digestion should produce fragments at 2697 and 4771bp instead of at 2579 and 4771bp.

7.3.1.3 Sequencing Data for the Wild-Type Vector Constructs

The Sanger sequencing data for the wild-type vector constructs are shown below, with the 3'-UTR inserts highlighted in yellow. “...” following a sequence indicates that the sequencing did not extend far enough to sequence the full insert, as was the case for the longer inserts.

7.3.1.3.1 BACH1 8

GTTTAACG GGTGTAGGGGAGGATATTGCTAGTATATTTTCAGTGGTTT
 GTATGTTCTCTCTGTCACTGACTTATTTGTAAGAGAAAATTAGTTGGACT
 TGTTTATTTTCTAGTAGCTTTTATAAGTACACTCAAGAATTTGTCAGGGA
 GAATAATTCTGATAGTGCATCCCATACTGCAAAGAATTTGTGTGTGTGT
 GTGTGTGTGTGTGTGTGTATGTGTATGTATACATATATATCTCTCCATAT
 AGGTATTTCTTTGATACTTGTAATTTTAAATTTTCAGCTTCACGATATAAAA
 TAATATAAGAACTTCTGGTTTACAAAATGTAATAATCTTAAGCCAATGGAA

CCCTTGATTTCCTACCTCAGTGTACACTCAACTATTGGTTGTATCAGTTT
GTGTATGTGCAAATGTCAAATAATCTTTTGCTTTAATTGCTACTGTACTT
GCTTTGAAAGATTACCTACTATTTTATGATAAAATGTAGTTGTCTCCAGA
GCTTAAATATAATTTGTAAAGCACTTGGTTTAAATTTCTCTCTACCTATAA
ACAGTTTAGCATTAAAGGGTTTCTATTAATGACACAGAATTATTGGCCAAG
TGTAATTTCTTAAAATTTAGCATTACTTTAAATAGCCAGCATGTAATACAA
GTA ACTACACTACCTCATATCTACATGATTTTCAAGTTGTAATGCAGATG
GACAGATAAAAAAGATTTTACGTTTGTCTTTTGGCCATAAGTGGGAAAGT
TTTCTGTATATTGCATAGCATTACACATTTATGCCTATTTTAAACATTA
TCTAAAGAAGTTTTTTCTAAGAAAATGTTTCAAGGCAATATTTTTTTGAG
GCTGCCGAAGACAAATGACAGGATTATGAGTATACAGTGTATGCCTTT
CCTTCATGCAGAATTTTGAATGTTTTCAGTTTGTATATTGCATATTCACA
TGATCATTGCTCACTATTTTATGAACTGGCCTTCTCAATGTTTGATGATT
TTTAAAAGCTGTTATG...

7.3.1.3.2 CDC25Aa 5

AGTTGTTTAACG GTGACATTTGGAGAGGGGGCCTGGGACTTCCATGCC
TTAAACCTACCTCCCACACTCCAAGGTTGGAGCCCAGGGCATCTTGCT
GGCTACGCCTCTTCTGTCCCTGT CGCTAGCCTCGAGTCTAGAGTCGAC
CTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAG
CTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG
GGCGGCCGCTTCGAGCAGACATGATAAGATAACATTGATGAGTTTGGA
CAAACCACA ACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG
TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAA
CAACAACAATTGCATTCATTTTATGTTTCAGGTT CAGGGGGAGATGTGG
GAGGTTTTTTAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAATT
TTAACAAAATATTAACGCTTACAATTCCTGATGCGGTATTTTCTCCTTAC
GCATCTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCA
TGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAG
GCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGT
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT
AGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCT
AACTCCGCCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCC

GCCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCC
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGC
CT

7.3.1.3.3 *CDC25Aa 6*

AACG**GTGACATTTGGAGAGGGGGCCTGGGACTTCCATGCCTTAAACCT**
ACCTCCCACACTCCCAAGGTTGGAGCCCAGGGCATCTTGCTGGCTACG
CCTCTTCTGTCCCTGTCGCTAGCCTCGAGTCTAGAGTCGACCTGCAGG
CATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTT
GGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGC
CGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAC
AACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTA
TTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAAC
AATTGCATTCATTTTATGTTTCAGGTTCCAGGGGGAGATGTGGGAGGTTT
TTTTAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAATTTTAACAA
AATATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCT
GTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCATGGCCT
GAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAA
AGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAG
GCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAG
CAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCC
GCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCA
TGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCC
TCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTAGGCT
TTTGCAAAGCTTGATTCTTCTGAC

7.3.1.3.4 *CDC25Ab 1*

TTAACG**CAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCTGTTAGACG**
TCCTCCGTCCATATCAGAACTGTGCCACAATGCAGTTCTGAGCACCGTG
TCAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCCGGGGCCTTATCG
GGCTCCAGCCATCTCATGAGGGGAGAGGAGACGGAGGGGAGTAGAGA
AGTTACACAGAAATGCTGCTGGCCAAATAGCAAAGACAACCTGGGAAG
GAAAGGTCTTTGTGGGATAATCCATATGTTTAATTTATTCAACTTCATCA
ATCACTTTATTTTATTTTTTTTTCTAACTCCTGGAGACTTATTTACTGCTT

CATTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTTATTATCCCAGGG
ACTACCTCGGCTTTTAATTTAAAAAAGAAAGTGGGTAAGAAAATGC
AAACCTGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCTGTCACCCCC
AGGAGGCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCCAAGAAGTGA
GGTCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTGACATAAGCT
AAATCTCCCAGACCCACCACTGGCTACCGATATCTATTTGGTGGGAGGT
GTGGCCCTGTTCTTCCTCACCCAGTTCCATGACATTGGCTGGTATAGG
AGCCACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTGGGCCACCCCC
GGCTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGGGAAGGGGGGT
GGGGGTAGGTGGGCTCCACAGGATGGGGGAGGAGCATGTCCACTGAG
TATCTTCCTTATGTTGCTGTGATATTGATAGCTTTTATTTTCTAATTTTAA
AAAATGGTCATATTATGAGTCAAAGAGTATCAAATCAGTGTTGGATGGA
CCACCCAAGGGTGAGGAAAGGGGCTGGAAGCCCTGGGCATTAGGAG...

7.3.1.3.5 CDC25Ab 2

TGATCTAGTTGTTTAACGCAGGGCATCTTGCTGGCTACGCCTCTTCTGT
CCCTGTTAGACGTCCTCCGTCCATATCAGAAGTGTGCCACAATGCAGTT
CTGAGCACCGTGTCAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCC
GGGGCCTTATCGGGCTCCAGCCATCTCATGAGGGGAGAGGAGACGGA
GGGGAGTAGAGAAGTTACACAGAAATGCTGCTGGCCAAATAGCAAAGA
CAACCTGGGAAGGAAAGGTCTTTGTGGGATAATCCATATGTTTAATTTAT
TCAACTTCATCAATCACTTTATTTTATTTTTTTTTCTAACTCCTGGAGACT
TATTTTACTGCTTCATTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTT
ATTATCCCAGGGACTACCTCGGCTTTTAATTTAAAAAAGAAAGTGA
GGTAAGAAAATGCAAACCTGTTATAAGTTATCGGACAGAAAGCTAGGTG
CTCTGTCACCCCCAGGAGGCGCTGTGGTACTGGGGCTGCTGCTATTTA
AGCCAAGAAGTGAAGTCCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTG
CCTGACATAAGCTAAATCTCCCAGACCCACCACTGGCTACCGATATCTA
TTTGGTGGGAGGTGTGGCCCTGTTCTTCCTCACCCAGTTCCATGACAT
TGGCTGGTATAGGAGCCACAGTCAGGAAAGCACTTGAGGCAGCATCTG
TTGGGCCACCCCCGGCTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCC
AGGGAAGGGGGGTGGGGGGTAGGTGGGCTCCACAGGATGGGGGAGG
AGCATGTCCACTGAGTATCTTCCTTATGTTGCTGTGATATTGATAGCTTT
TATTTTCTAATTTTAAAAAATGGTCATATTATGAGTCAAAGAGTATCAA

TCAGTGTGGATGGACCACCCAAGGGTGAGGAAAGGGGCTGGAAGCC
CTGGGCATTAGG[AAAAAG]*GGA... *Apparent sequencing error.

7.3.1.3.6 *CDC25Ab 4*

TAACGCAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCTGTTAGACGT
CCTCCGTCCATATCAGAACTGTGCCACAATGCAGTTCTGAGCACCGTGT
CAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCCGGGGCCTTATCGG
GCTCCAGCCATCTCATGAGGGGAGAGGAGACGGAGGGGAGTAGAGAA
GTTACACAGAAATGCTGCTGGCCAAATAGCAAAGACAACCTGGGAAGG
AAAGGTCTTTGTGGGATAATCCATATGTTTAATTTATTCAACTTCATCAAT
CACTTTATTTATTTTTTTTTCTAACTCCTGGAGACTTATTTACTGCTTCA
TTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTTATTATCCAGGGACT
ACCTCGGCTTTTAATTTAAAAAAAAAAAAAGAAGTGGGTAAGAAAATGCAA
ACCTGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCTGTCACCCCCAG
GAGGCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCCAAGAAGTGGAGG
TCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTGACATAAGCTAA
ATCTCCCAGACCCACCACTGGCTACCGATATCTATTTGGTGGGAGGTGT
GGCCCTGTTCTTCCTCACCCCAGTTCCATGACATTGGCTGGTATAGGAG
CCACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTGGGCCACCCCCGG
CTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGGGAAGGGGGGTGG
GGGTAGGTGGGCTCCACAGGATGGGGGAGGAGCATGTCCACTGAGTA
TCTTCCTTATGTTGCTGTGATATTGATAGCTTTTATTTTCTAATTTTTAAAA
AATGGTCATATTATGAGTCAAAGAGTATCAAATCAGTGTGGATGGACC
ACCAAGGGTGAGGAGAGGGGCTGGAAGCCCTGGGCATTAGGAGAAG
GGAG...

7.3.1.3.7 *HMGA2a 3*

TCAGTTGTTTAACGGTTCGATTTCTACCTCAGCAGCAGTTGGATCTTTTG
AAGGGAGAAGACACTGCAGTGACCACTTATTCTGTATTGCCATGGTCTT
TCCACTTTCATCTGGGGTGGGGTGGGGTGGGGTGGGG[GAAAAAA]*

*Apparent sequencing error.

7.3.1.3.8 *HMGA2a 4*

TCAGTTGTTTAACGGTTCGATTTCTACCTCAGCAGCAGTTGGATCTTTTG
AAGGGAGAAGACACTGCAGTGACCACTTATTCTGTATTGCCATGGTCTT

TCCACTTTCATCTGGGGTGGGGTGGGGTGGGGTGGGGGA[*TGAGAGT*]*

*Apparent sequencing error.

7.3.1.3.9 *HMGA2b 1*

TTGTTTAACGGCTGCTTCAGGGAGGTAGTTTCAAAGGCCACATACCTCT
CTGAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTATG
GAGAGAATTA~~AA~~ACTCAACATTACTGTAACTGTGCGTTAAATAAGCAAA
TAAACAGTGGCTCATA~~AA~~AATA~~AA~~AAGTCGCATTCCATATCTTTGGATGGG
CCTTTTAGAAACCTCATTGGCCAGCTCATA~~AA~~AATGGAAGCAATTGCTCAT
GTTGGCCAAACATGGTGCACCGAGTGATTCCATCTCTGGTAAAGTTAC
ACTTTTATTTCTGTATGTTGTACAATCAAACACACTACTACCTCTTAAG
TCCCAGTATACCTCATT~~TTT~~CATACTGAAAAAAAAAAGCTTGTGGCCAATG
GAACAGTAAGAACATCATA~~AA~~AATTTTATATATATAGTTTATTTTGTGGG
AGATAAATTTTATAGGACTGTTCTTTGCTGTTGTTGGTCGCCGCTAGCCT
CGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAAC
AAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAA
CTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATGATAAGATA
CATTGATGAGTTTGGACAAACCACA~~ACT~~AGAATGCAGTGAAAAAATGC
TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGC
TGCAATAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTTTCAG

7.3.1.3.10 *HMGA2b 2*

TGTTTACGGCTGCTTCAGGGAGGTAGTTTCAAAGGCCACATACCTCTCT
GAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTATGGA
GAGAATTA~~AA~~ACTCAACATTACTGTAACTGTGCGTTAAATAAGCAAATA
AACAGTGGCTCATA~~AA~~AATA~~AA~~AAGTCGCATTCCATATCTTTGGATGGGC
CTTTTAGAAACCTCATTGGCCAGCTCATA~~AA~~AATGGAAGCAATTGCTCAT
GTTGGCCAAACATGGTGCACCGAGTGATTCCATCTCTGGTAAAGTTAC
ACTTTTATTTCTGTATGTTGTACAATCAAACACACTACTACCTCTTAAG
TCCCAGTATACCTCATT~~TTT~~CATACTGAAAAAAAAAAGCTTGTGGCCAATG
GAACAGTAAGAACATCATA~~AA~~AATTTTATATATATAGTTTATTTTGTGGG
AGATAAATTTTATAGGACTGTTCTTTGCTGTTGTTGGTCGCCGCTAGCCT
CGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAAC
AAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAA
CTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATGATAAGATA

CATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGC
TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGC
TGCAATAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTT
CAGGGGGAGATGTGGGAAGGTTTTTTAAGCAAGTAAACCTCTACAAA
TGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTCCTGAT
GCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGC
C

7.3.1.3.11 *HMGA2b 3*

AGTTGTTTAACG **GCTGCTTCAGGGAGGTAGTTTCAAAGGCCACATACCT**
CTCTGAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTAT
GGAGAGAATTAAACTCAACATTACTGTAACTGTGCGTTAAATAAGCAA
ATAAACAGTGGCTCATAAAAATAAAAGTCGCATTCCATATCTTTGGATGG
GCCTTTTAGAACCTCATTGGCCAGCTCATAAAATGGAAGCAATTGCTC
ATGTTGGCCAAACATGGTGCACCGAGTGATTTCCATCTCTGGTAAAGTT
ACACTTTTATTTCTGTATGTTGTACAATCAAAACACACTACTACCTCTTA
AGTCCCAGTATACCTCATTTCATACTGAAAAAAAAAGCTTGTGGCCAA
TGGAACAGTAAGAACATCATAAAATTTTATATATATAGTTTATTTTTGTG
GGAGATAAATTTTATAGGACTGTTCTTTGCTGTTGTTGGTCGCCGCTAG
CCTCGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCT
AACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATGATAAG
ATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAAT
GCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAA
GCTGCAATAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAG
GTTTCAGGGGGAGATGTGGGAGGTTTTTTAAGCAAGTAAACCTCTACA
AATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTCCTGA
TGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACG
CGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTG

7.3.1.3.12 *HMGA2b 4*

CTAGTTGTTTAACG **GCTGCTTCAGGGAGGTAGTTTCAAAGGCCACATAC**
CTCTCTGAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCT
ATGGAGAGAATTAAACTCAACATTACTGTAACTGTGCGTTAAATAAGC
AAATAAACAGTGGCTCATAAAAATAAAAGTCGCATTCCATATCTTTGGAT

GGGCCTTTTAGAAACCTCATTGGCCAGCTCATAAAATGGAAGCAATTGC
TCATGTTGGCCAAACATGGTGCACCGAGTGATTTCCATCTCTGGTAAAG
TTACACTTTTATTTCTGTATGTTGTACAATCAAAACACACTACTACCTCT
TAAGTCCCAGTATACCTCATTTCATACTGAAAAAAAAAAGCTTGTGGCC
AATGGAACAGTAAGAACATCATAAAATTTTTATATATATAGTTTATTTTTG
TGGGAGATAAATTTTATAGGACTGTTCTTTGCTGTTGTTGGTCGCCGCT
AGCCTCGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGC
AATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATGATA
AGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAA
AATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA
TAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTC
AGGTTCAAGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAACCTCTA
CAAATGTGGTAAAATCGAATTTAACAAAATATTAACGCTTACAATTTCT
GATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATA
CGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTGAAAGA

7.3.1.3.13 KRAS 2

GTATCTAGTTGTTTAACGGGTCTCTGTGCCAGCTCTATAATTGTTTTGCT
ACGATTCCACTGAAACTCTTCGATCAAGCTACTTTATGTAAATCACTTCA
TTGTTTTAAAGGAATAAACTTGATTATATTGTTTTTTTATTTGGCATAACT
GTGATTCTTTTGGGACAATTAAGGTGTATGTCAGATAT
TCATATTGACCCAAATGTGTAATATTCCAGTTTTCTCTGCATAAGTAATTA
AAATATACTTAAAAATTAATAGTTTTATCTGGGTACAAATAAACAGGTGC
CTGAACTAGTTCACAGACAAGGAACTTCTATGTAAAAATCACTATGATT
TCTGAATTGCTATGTGAAACTACAGATCTTTGGAACACTGTTTAGGTAGG
GTGTTAAGACTTACACAGTACCTCGTTTCTACACAGAGAAAGAAATGGC
CATACTTCAGGAACTGCAGTGCTTATGAGGGGATATTTAGGCCTCTTGA
ATTTTTGATGTAGATGGGCATTTTTTTAAGGTAGTGGTAATTACCTTTAT
GTGAACTTTGAATGGTTTAAACAAAAGATTTGTTTTTGTAGAGATTTAAA
GGGGGAGAATTCTAGAAATAAATGTTACCTAATTATTACAGCCTTAAAGA
TAAAAATCCTTGTTGAAGTTTTTTAAAAAAAAGCTAAATTACATAGACTTA
GGCATTAAACATGTTTGTGGAAGAATATAGCAGACGTATATTGTATCATT
GAGTGAATGTTCCCAAGTAGGCA CGCTAGCCTCGAGTCTAGAGTCGAC

CTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAG
CTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG
GGGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGA
CAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG
GGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAC

7.3.1.3.14 KRAS 3

GTTTAACGGGTCTCTGTGCCAGCTCTATAATTGTTTTGCTACGATTCCAC
TGAAACTCTTCGATCAAGCTACTTTATGTAAATCACTTCATTGTTTTAAAG
GAATAAACTTGATTATATTGTTTTTTTATTTGGCATAACTGTGATTCTTTT
GGGACAATTACTGTACACATTAAGGTGTATGTCAGATATTCATATTGACC
CAAATGTGTAATATTCCAGTTTTCTCTGCATAAGTAATTAATAATACTTA
AAAATTAATAGTTTTATCTGGGTACAAATAAACAGGTGCCTGAACTAGTT
CACAGACAAGGAACTTCTATGTAAAAATCACTATGATTTCTGAATTGCT
ATGTGAAACTACAGATCTTTGGAACACTGTTTAGGTAGGGTGTAAAGAC
TTACACAGTACCTCGTTTCTACACAGAGAAAGAAATGGCCATACTTCAG
GAACTGCAGTGCTTATGAGGGGATATTTAGGCCTCTTGAATTTTTGATG
TAGATGGGCATTTTTTAAGGTAGTGGTAATTACCTTTATGTGAACTTTG
AATGGTTTAACAAAAGATTTGTTTTTG TAGAGATTTTAAAGGGGGAGAAT
TCTAGAAATAAATGTTACCTAATTATTACAGCCTTAAAGATAAAAAATCCTT
GTTGAAGTTTTTTAAAAAAAAGCTAAATTACATAGACTTAGGCATTAACAT
GTTTGTGGAAGAATATAGCAGACGTATATTGTATCATTGAGTGAATGTT
CCCAAGTAGGCACGCTAGCCTCGAGTCTAGAGTCGACCTGCAGGCATG
CAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCT
GCTGCCACCGCTGAACAATAACTAGCATAACCCCTTGGGGCGGCCGCT
TCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC
AGAATGCAGTGAAAAAATGCTTTATTTGGGAAATTTGGGATGCT

7.3.1.3.15 KRAS 4

GTTGTTTAAACGGGTCTCTGTGCCAGCTCTATAATTGTTTTGCTACGATT
CCACTGAAACTCTTCGATCAAGCTACTTTATGTAAATCACTTCATTGTTTT
AAAGGAATAAACTTGATTATATTGTTTTTTTATTTGGCATAACTGTGATTC
TTTTGGGACAATTACTGTACACATTAAGGTGTATGTCAGATATTCATATT

GACCCAAATGTGTAATATTCCAGTTTTCTCTGCATAAGTAATTAATAATA
CTTAAAAATTAATAGTTTTATCTGGGTACAAATAAACAGGTGCCTGA
AGTTCACAGACAAGGAACTTCTATGTAAAAATCACTATGATTTCTGA
TGCTATGTGAACTACAGATCTTTGGAACACTGTTTAGGTAGGGTGTTA
AGACTTACACAGTACCTCGTTTTCTACACAGAGAAAGAAATGGCCATA
TCAGGAACTGCAGTGCTTATGAGGGGATATTTAGGCCTCTTGAATTTT
GATGTAGATGGGCATTTTTTTAAGGTAGTGGTAATTACCTTTATGTGAAC
TTTGAATGGTTTAAACAAAAGATTTGTTTTGTAGAGATTTTAAAGGGGGA
GAATTCTAGAAATAAATGTTACCTAATTATTACAGCCTTAAAGATAAAAAT
CCTTGTTGAAGTTTTTTAAAAAAAAGCTAAATTACATAGACTTAGGCATTA
ACATGTTTGTGGAAGAATATAGCAGACGTATATTGTATCATTGAGTGAA
TGTTCCAAGTAGGCA CGCTAGCCTCGAGTCTAGAGTCGACCTGCAGG
CATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTT
GGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGC
CGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAC
AACTAGAATGCAGTGAAAAAATGCTTTATTTGGGAAATTTGGGATG

7.3.1.3.16 MYC 1

TTAACG AATGTCCTGAGCAATCACCTATGAACTTGTTTCAAATGCATGAT
CAAATGCAACCTCACAACCTTGGCTGAGTCTTGAGACTGAAAGATTTAG
CCATAATGTAACTGCCTC CGCTAGCCTCGAGTCTAGAGTCGACCTGCA
GGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCG
GCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACC
ACA ACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGC
TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAA
CAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTT
TTTTTAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAATTTTAA
AAATATTAACGCTTACAATTCCTGATGCGGTATTTTCTCCTTACGCATC
TGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCATGGCC
TGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGA
AAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCA
GGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCA
GCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATG

CAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTC
CGCCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCC
ATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTAGG
CTTTTGCAAAAAGCTTGATTCTTCTGACCC

7.3.1.3.18 MYC 2

GTTGTTTAACG AATGTCCTGAGCAATCACCTATGAACTTGTTTCAAATGC
ATGATCAAATGCAACCTCACAACCTTGGCTGAGTCTTGAGACTGAAAGA
TTTAGCCATAATGTAAACTGCCTC CGCTAGCCTCGAGTCTAGAGTCGAC
CTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAAG
CTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG
GGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGA
CAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG
TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAA
CAACAACAATTGCATTCATTTTATGTTTCAGGTT CAGGGGGAGATGTGG
GAGGTTTTTTTAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAATT
TTAACAAAATATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTAC
GCATCTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCACCA
TGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAG
GCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGT
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT
AGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCT
AACTCCGCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCC
GCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCC
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGC
CTA

7.3.1.3.19 MYC 3

ACTTGTTTAACG AATGTCCTGAGCAATCACCTATGAACTTGTTTCAAATG
CATGATCAAATGCAACCTCACAACCTTGGCTGAGTCTTGAGACTGAAAG
ATTTAGCCATAATGTAAACTGCCTC CGCTAGCCTCGAGTCTAGAGTCGA
CCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGGAA
GCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTT

GGGGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGG
ACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTT
AACAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGATGT
GGGAGGTTTTTTTTAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGAA
TTTTAACAAAATATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTT
ACGCATCTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCAC
CATGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTG
AGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAA
GTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAA
TTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCAGGCAG
AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCG
CCT

7.3.1.3.20 “Let-7f-5p Antisense Repeats” 1

TCATCTAGTTGTTTAACGAACTATAACAATCTACTACCTCAAACCTATAACAAT
CTACTACCTCAAACCTATAACAATCTACTACCTCAAACCTATAACAATCTACTA
CCTCACGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTG
ATCCGGCTGCTAACAAAGCCC GAAAGGAAGCTGAGTTGGCTGCTGCCA
CCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCA
GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGC
AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG
TAACCATTATAAGCTGCAATAACAAGTTAACAAACAACAATTGCATTCAT
TTTATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTTAAGCAAGT
AAAACCTCTACAAATGTGGTAAAATCGAATTTTAAACAAAATATTAACGCT
TACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTT
ACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCT
GAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTG
TGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCAGGCTCCCAGCA
GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTG
GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATC
TCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC

CCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATT
TTTTTTATTTATGCAGA

7.3.1.3.21 *“Let-7f-5p Antisense Repeats” 2*

TATCTAGTTGTTTAACGAACTATACAATCTACTACCTCAAACCTATACAATC
TACTACCTCAAACCTATACAATCTACTACCTCAAACCTATACAAACCTACTA
CCTCACGCTAGCCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCT
GATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCC
ACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGC
AGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATG
CAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT
GTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCA
TTTTATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAG
TAAACCTCTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGC
TTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTT
CACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTC
TGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCT
GTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGT
GGAAAGTCCCCAGGCTCCCCAGCAGCAGAAGTATGCAAAGCATGCATC
TCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC
CCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATT
TTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGGCCTCTGAGCTATTC
C

7.3.1.3.22 *“Let-7f-5p Antisense Repeats” 4*

TATCTAGTTGTTTAACGAACTATACAATCTACTACCTCAAACCTATACAATC
TACTACCTCAAACCTATACAATCTACTACCCTCAAACCTATACAATCTACTA
CCTCACGCTAGCCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTG
ATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCA
CCGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCA
GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGC
AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG
TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAT
TTTATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGT

AAAACCTCTACAAATGTGGTAAAATCGAATTTTAAACAAAATATTAACGCT
TACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTT
ACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCT
GAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTG
TGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCA
GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTG
GAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATC
TCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC
CCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATT
TTTTTATTTATGCAGAGGC

7.3.1.3.23 “Let-7f-5p Antisense Repeats” 5

TATCTAGTTGTTTAAACGAACCTATAACAATCTACTACCTCAAACCTATAACAATC
TACTACCTCAAACCTATAACAATCTACTACCTCAAACCTATAACAATCTACTAC
CTCACGCTAGCCTCGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGA
TCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAG
ACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCA
GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGT
AACCATTAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT
TTATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTTAAGCAAGTA
AAACCTCTACAAATGTGGTAAAATCGAATTTTAAACAAAATATTAACGCTT
ACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCA
CACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTG
AAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGT
GGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAG
GCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGG
AAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCT
CAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC
CCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATT
TTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCC
A

7.3.1.3.24 “Let-7f-5p Antisense Repeats” 9

ATCTAGTTGTTTAAACGAAGCTATACAATCTACTACCTCAAACCTATACAATCT
ACTACCTCAAACCTATACAATCTACTACCTCAAACCTATACAATCTACTACC
TCACGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACC
GCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGA
CATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAG
TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA
ACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTT
TATGTTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAA
AACCTCTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTA
CAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTAC
ACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTGA
AAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTG
GAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG
CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGCAGAAGTATGCAAAGCATGCATCTCA
ATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCC
TAACTCCGCCCAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTT
TTTATTTATGCAGAGGCCGAGGCCCGCCTCGGCCTCTGAGCTATTCC

7.3.1.3.25 “Let-7f-5p Antisense Repeats” 10

CTAGTTGTTTACGAAGCTATACAATCTACTACCTCAAACCTATACAATCTACT
ACCTCAAACCTATACAATCTACTACCTCAAACCTATACAATCTACTACCTCA
CGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTGATCCG
GCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCT
GAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACAT
GATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGA
AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACC
ATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG
TTTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAAAACC
TCTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATT
TCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTACACCG
CATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTGAAAGA

GGAAGTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGA
AGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGT
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT
AGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAA
CTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATTTTTTT
ATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAG
TAGTGAGG

7.3.1.3.26 “Let-7f-5p Antisense Repeats” 11

CTAGTTGTTTACGAACTATAACAATCTACTACCTCAAACCTATAACAATCTACT
ACCTCAAACCTATAACAATCTACTACCTCAAACCTATAACAATCTACTACCTCA
CGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTGATCCG
GCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCT
GAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACAT
GATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGA
AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACC
ATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG
TTTCAGGTTCCAGGGGGAGATGTGGGAGGTTTTTTTTAAGCAAGTAAACC
TCTACAAATGTGGTAAAATCGAATTTTAAACAAAATATTAACGCTTACAATT
TCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTACACCG
CATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTGAAAGA
GGAAGTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGA
AGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGT
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT
AGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAA
CTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATTTTTTT
TATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAG
TAGTGAGGAGGCTTTTTTTGGA

7.3.1.3.27 “Let-7f-5p Antisense Repeats” 13

ATCTAGTTGTTTAAACGAACTATAACAATCTACTACCTCAAACCTATAACAATCT
ACTACCTCAAACCTATAACAATCTACTACCTCAAACCTATAACAATCTACTACC
TCACGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTGAT

CCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACC
GCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGA
CATGATAAGATACATTGATGAGTTTGGACAAACCACA ACTAGAATGCAG
TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA
ACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTT
TATGTTTCAGGTT CAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTAA
AACCTCTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTA
CAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTTAC
ACCGCATAACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTGA
AAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTG
GAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG
CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTC
AATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCC
CTAACTCCGCCAGTTCGCCCATCTCCGCCCATGCTGACTAATTTTTT
TATTTATG

7.3.1.3.28 *“Let-7f-5p Antisense Repeats” 15*

TTTCTAGTTGTTTAACGAACTATACAATCTACTACCTCAA ACTATACAATC
TACTACCTCAA ACTATACAATCTACTACCTCAA ACTATACAATCTACTAC
CTCACGCTAGCCTCGAGTC TAGAGTCGACCTGCAGGCATGCAAGCTGA
TCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAG
ACATGATAAGATACATTGATGAGTTTGGACAAACCACA ACTAGAATGCA
GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGT
AACATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT
TTATGTTTCAGGTT CAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGTA
AAACCTCTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTT
ACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCA
CACCGCATAACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCTG
AAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGT
GGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAG
GCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGG
AAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCT

CAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCC
 CCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATT
 TTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCC

7.3.1.4 Primer Design for PCR Mutagenesis of Predicted Target Loci

As described in the Methods, primers were designed to mutate the predicted miRNA-targeted loci, particularly at seed regions. The full sequences of the primers used are listed in Table 19.

Table 19: A list of the full sequences of forward and reverse primers designed for the PCR mutagenesis of predicted miRNA-targeted loci in 3'-UTRs.

Primer	Sequence (5' to 3' Order)
BACH1-i-mut-FW	CCAATGGAACCCTTGATTTGATATCCAGTGTACACCCAAC TATTGGTTGTATCAG
BACH1-i-mut-RV	CTGATACAACCAATAGTTGGGTGTACACTGGATATCGAAAT CAAGGGTTCCATTGG
BACH1-ii-mut-FW	GCCAGCATGTAATACAAGTAACTACAGGATCCCATATCTAC ATGATTTTCAAGTTGTAATGCAGATGG
BACH1-ii-mut-RV	CCATCTGCATTACAACCTGAAAATCATGTAGATATGGGATC CTGTAGTTACTTGTATTACATGCTGGC
CDC25Aa-mut-FW	CTTCCATGCCTTAAACCGATATCCCACACTCCCAAGGTTG
CDC25Aa-mut-RV	CAACCTTGGGAGTGTGGGATATCGGTTTAAGGCATGGAAG
CDC25Ab-i-mut-FW	CTAGGTAGGGTTTTATTATCCAGGGACGGATCCGGCTTTT AATTTAAAAAAAAAAAAAGAAGTGGG
CDC25Ab-i-mut-RV	CCCCTTCTTTTTTTTTTAAATTAAGCCGGATCCGTCC CTGGGATAATAAAACCCTACCTAG
CDC25Ab-ii-mut-FW	CAGGAGACAGTCTCTCCCTCTCAATTGAACTTCTTGGGGT GGAATC
CDC25Ab-ii-mut-RV	GATTCCCACCCCAAGAAGTTCAATTGAGAGGGGAGAGACT GTCTCCTG
HMGA2a-i-mut-FW	TAATTCTAGTTGTTTAAACGGTTCGATTTCCGGATCCAGCAG CAGTTGGATCTTTTGAAG
HMGA2a-i-mut-RV	CTTCAAAGATCCAACCTGCTGCTGGATCCGAAATCGAACCG TTTAAACAACCTAGAATTA
HMGA2a-ii-mut-FW	CCATGAAAGACCTGAATACCACTGAATTCAAATTAAGCATAT GTGTTACTTCAAG
HMGA2a-ii-mut-RV	CTTGAAGTAACACATATGCTTAATTTGAATTCAGTGGTATTC AGGTCTTTCATGG
HMGA2a-iii-mut-FW	GCCAGAGATGCAATATATCCCCACTACTCAATACCAATTGT GAATGTTACAACGAATTTACAGTCTAG
HMGA2a-iii-mut-RV	CTAGACTGTAAATTCGTTGTAACATTCACAATTGGTATTGAG TAGTGGGGATATATTGCATCTCTGGC
HMGA2a-iv-mut-FW	GATTTGACTGTAATATTTAAATATTACCCTCCAAGTCTGACG CGTAAATGAATCTTTAAGGAGATGGAC

(Table 19 Continued)

Primer	Sequence (5' to 3' Order)
HMGA2a-iv-mut-RV	GTCCATCTCCTTAAAGAATTCATTTACGCGTCAGACTTGGA GGGTAATATTTAAATATTACAGTCAAATC
HMGA2a-v-mut-FW	GGAGATGGACTAATTGACTTGCAAAGACCCATATGCAGACT TCAAAGGAATGAACTTGTTAC
HMGA2a-v-mut-RV	GTAACAAGTTCATTCCTTTTGAAGTCTGCATATGGGTCTTTG CAAGTCAATTAGTCCATCTCC
HMGA2b-vi-mut-FW	GAGGTAGTTTCAAAGGCCACAGATATCTCTGAGACTGGCA GATCGCTC
HMGA2b-vi-mut-RV	GAGCGATCTGCCAGTCTCAGAGATATCTGTGGCCTTTGAAA CTACCTC
HMGA2b-vii-mut-FW	CCTGTATGTTGTACAATCAAACACACTACGGATCCTTAAG TCCCAGTATACCTCATTTTTTCATAC
HMGA2b-vii-mut-RV	GTATGAAAAATGAGGTATACTGGGACTTAAGGATCCGTAGT GTGTTTTGATTGTACAACATACAGG
HMGA2b-viii-mut-FW	GTACAATCAAACACACTACTACCTCTTAAGTCCCAGTAAG ATCTATTTTTCATACTGAAAAAAAAAAGCTTGTGGCCAATGG
HMGA2b-viii-mut-RV	CCATTGGCCACAAGCTTTTTTTTTTTCAGTATGAAAAATAGATC TACTGGGACTTAAGAGGTAGTAGTGTGTTTTGATTGTAC
KRAS-mut-FW	GGGTGTTAAGACTTACACAGGATATCGTTTCTACACAGAGA AAGAAATG
KRAS-mut-RV	CATTTCTTTCTCTGTGTAGAAACGATATCCTGTGTAAGTCTT AACACCC
MYC-mut-FW	GAGACTGAAAGATTTAGCCATAATGTAAAGGATCCCAAATT GGACTTTGGGCATAAAAGAAC
MYC-mut-RV	GTTCTTTTATGCCCAAAGTCCAATTTGGGATCCTTTACATTA TGGCTAAATCTTTCAGTCTC

7.3.1.5 Checking Mutant Vectors for Successful Binding Site Mutations by Digestion

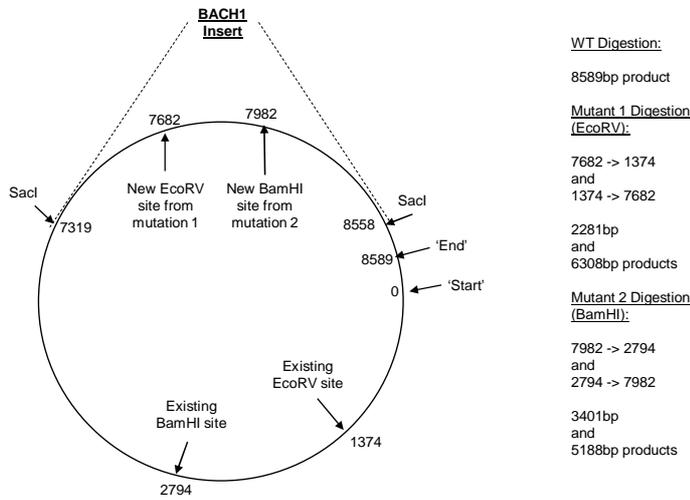


Figure 105: A diagram (not to scale) illustrating the loci of the existing restriction sites in the BACH1 construct, and those of new restriction sites created by successful seed region mutagenesis

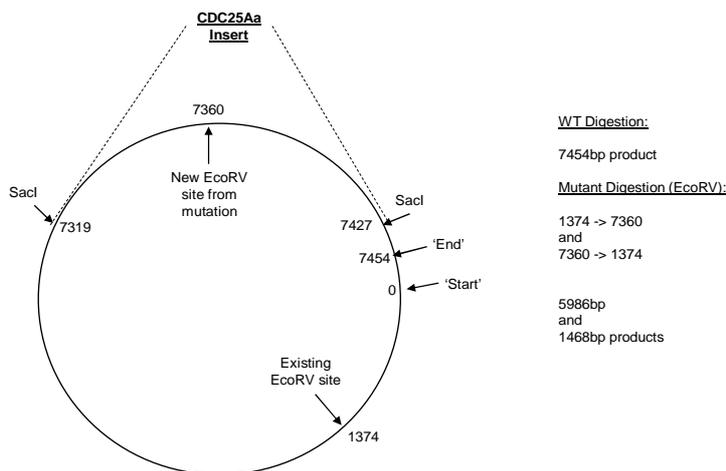


Figure 106: A diagram (not to scale) illustrating the loci of the existing restriction sites in the CDC25Aa construct, and those of new restriction sites created by successful seed region mutagenesis

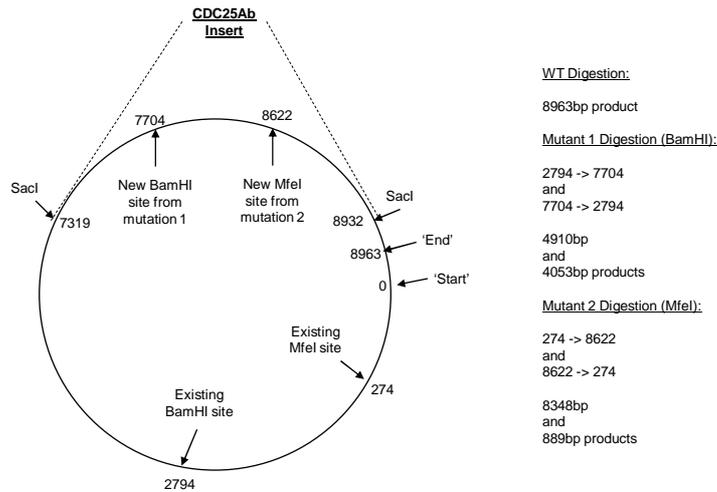


Figure 107: A diagram (not to scale) illustrating the loci of the existing restriction sites in the CDC25Ab construct, and those of new restriction sites created by successful seed region mutagenesis

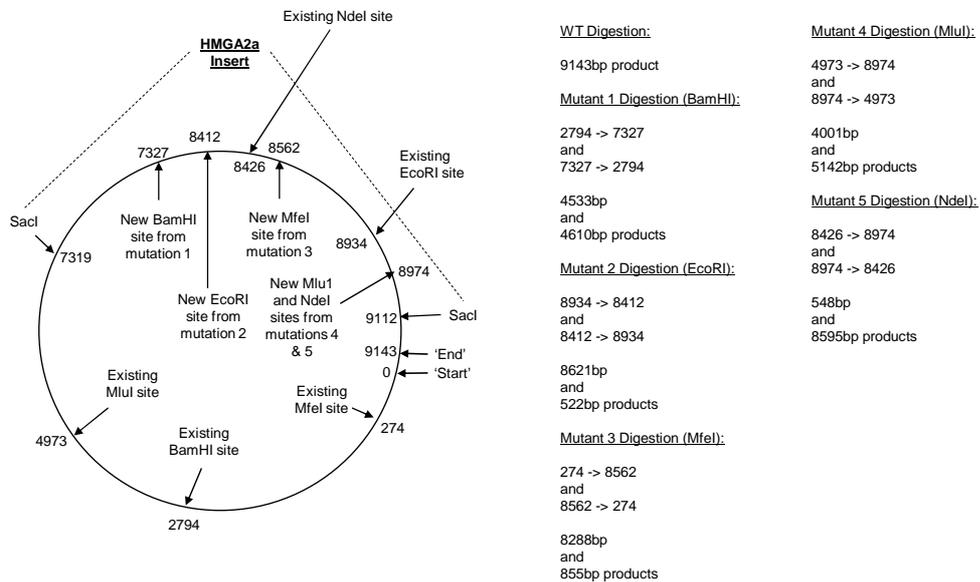


Figure 108: A diagram (not to scale) illustrating the loci of the existing restriction sites in the HMGA2a construct, and those of new restriction sites created by successful seed region mutagenesis

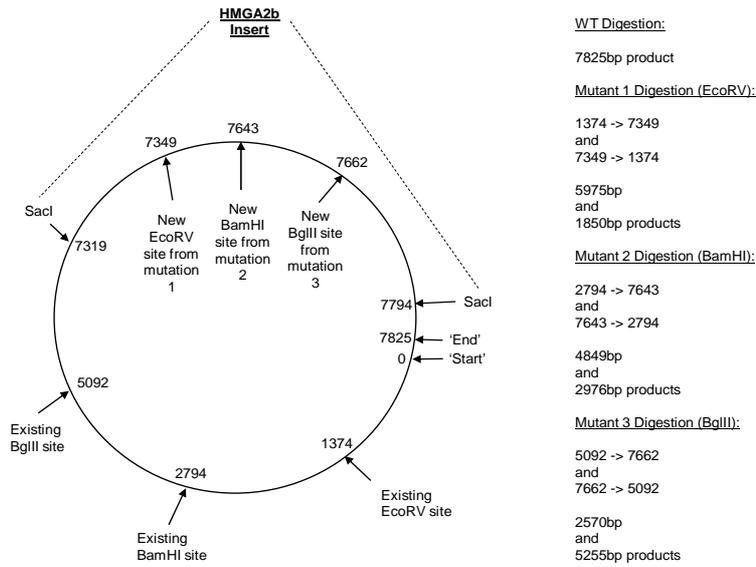


Figure 109: A diagram (not to scale) illustrating the loci of the existing restriction sites in the HMGA2b construct, and those of new restriction sites created by successful seed region mutagenesis

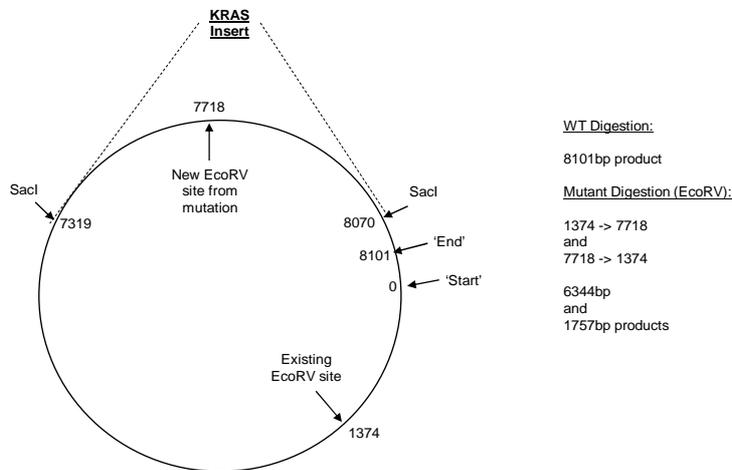


Figure 110: A diagram (not to scale) illustrating the loci of the existing restriction sites in the KRAS construct, and those of new restriction sites created by successful seed region mutagenesis

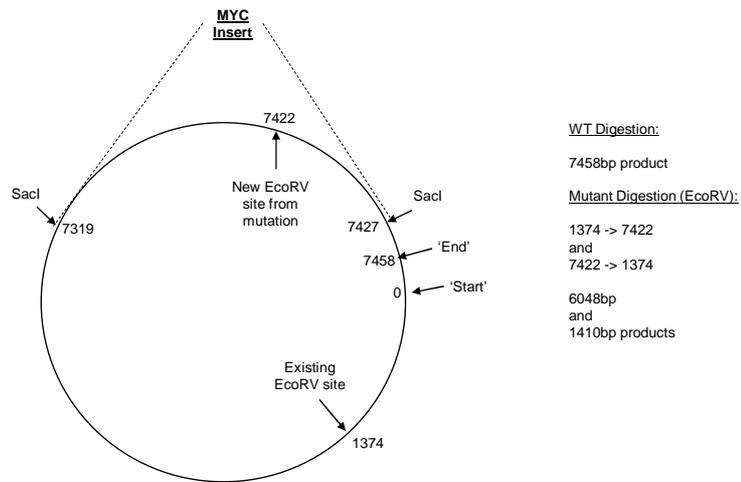


Figure 111: A diagram (not to scale) illustrating the loci of the existing restriction sites in the MYC construct, and those of new restriction sites created by successful seed region mutagenesis

Figure 105, Figure 106, Figure 107, Figure 108, Figure 109, Figure 110 and Figure 111 all show the existing restriction sites present in the wild-type constructs, and the loci of those newly created by successful seed region mutation. The expected products following digestion are listed alongside each diagram.

7.3.1.5.1 Agarose Electrophoresis of Digestion Products

7.3.1.5.1.1 First-Site Mutagenesis: Bach1, CDC25Ab, HMGA2a, HMGA2b and KRAS

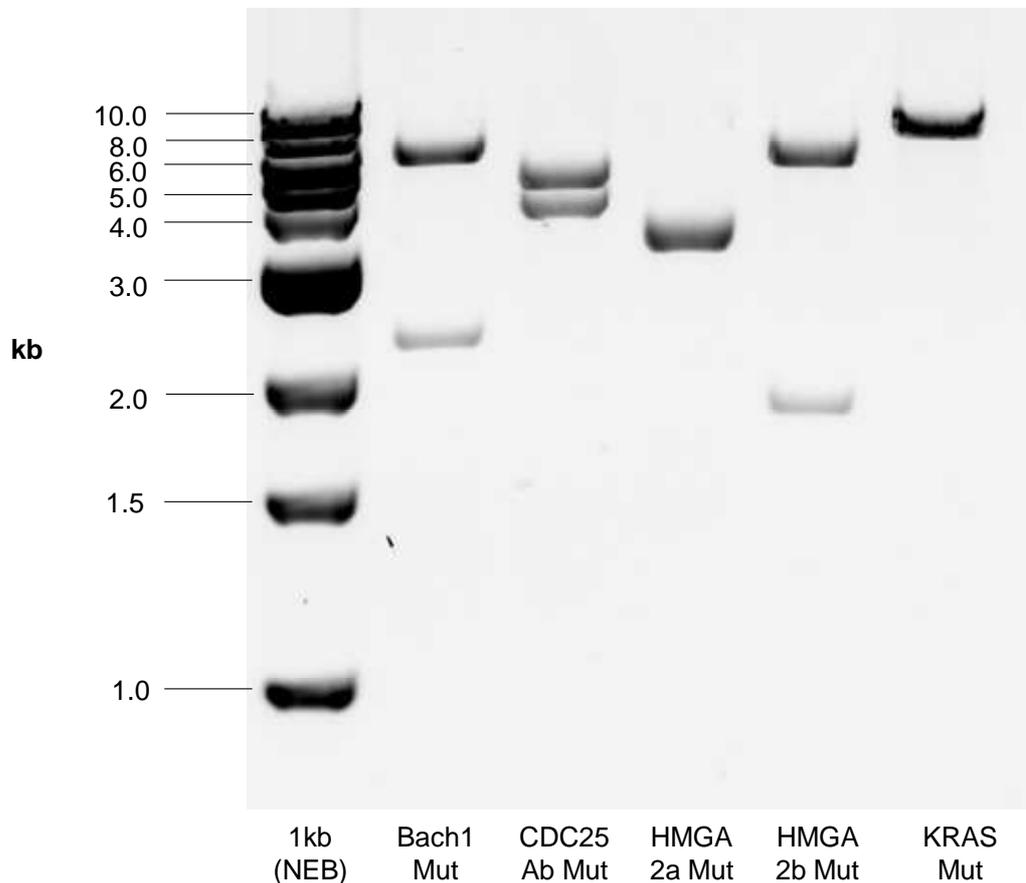


Figure 112: An ethidium bromide-based image of a 1% (w/v) agarose gel into which the EcoRV digestion products of BACH1 and HMGA2b first-site mutants, and the BamHI digestion products of CDC25Ab, HMGA2a and KRAS first-site mutants were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 80V for 2h.

It is clear from the gel image in Figure 112 that the first binding site of the BACH1 construct was successfully mutagenized, creating a new EcoRV restriction site, resulting in the expected 2281 and 6308bp products, as described in Figure 105. Mutageneses of the first binding sites of CDC25Ab and HMGA2b were also successful, as illustrated by the presence of

expected 4053 and 4910bp, and 1850 and 5975bp bands, respectively (see Figure 107 and Figure 109 for detail about expected products). For HMGA2a, there should be 4533 and 4610bp products (see Figure 108), which may not be separate from one another as discrete bands due to their proximity. However, the band in this lane according to Figure 112 actually appears between 3.0 and 4.0 kb, thus lower than expected.

7.3.1.5.1.2 Second-Site Mutagenesis: Bach1, CDC25Ab and HMGA2b

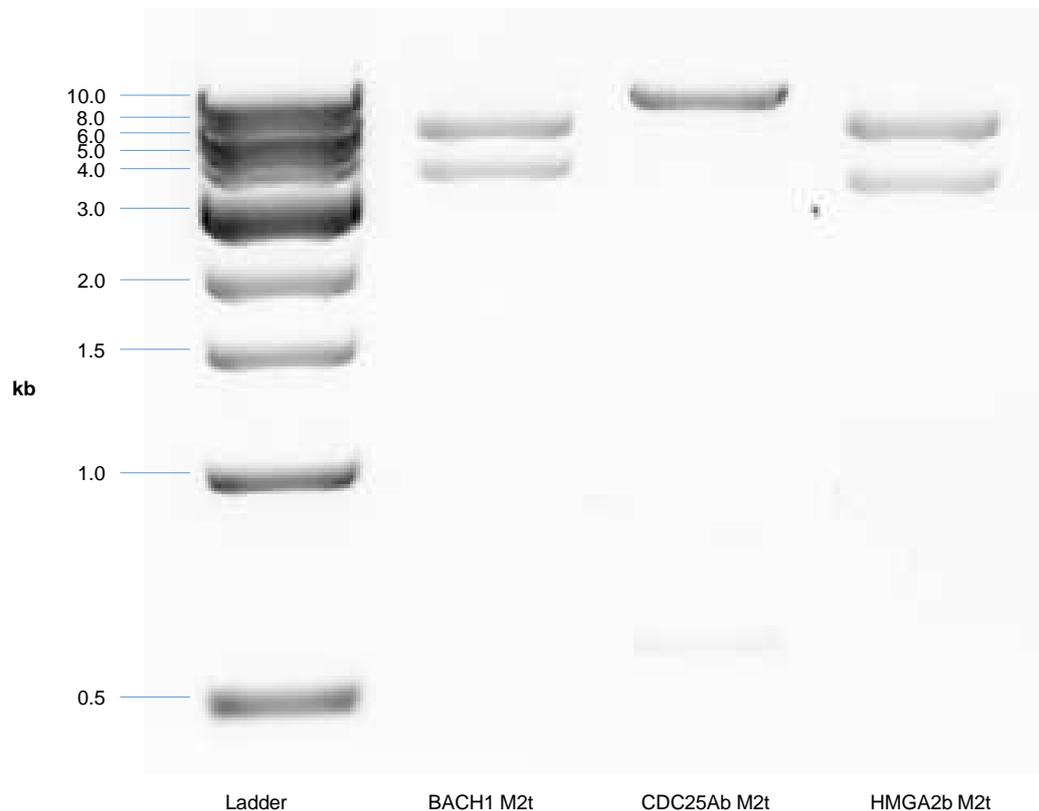


Figure 113: An ethidium bromide-based image of a 1% (w/v) agarose gel into which the BamHI digestion products of BACH1 and HMGA2b double mutants, and the MfeI digestion product of CDC25Ab double mutant were loaded, alongside a 1kb molecular weight marker (NEB), and electrophoresed at 60V for 1h.

7.3.1.6 Sequencing Data for the Mutant Vectors

7.3.1.6.1 BACH1 – 1st Seed Mutation

The FW and antiparallel of RV sequences were combined:

CTAGTTGTTAACGGGTGTAGGGGGAGGATATTGCTAGTATATTTTCAG
TGGTTTGTATGTTCTCTCTGTCACTGACTTATTTGTAAGAGAAAATTAGT
TGGACTTGTTTATTTCTAGTAGCTTTTATAAGTACACTCAAGAATTTGTC
AGGGAGAATAATTCTGATAGTGCATCCCATACTGCAAAGAATTTGTGT
GTGTGTGTGTGTGTGTGTGTGTGTATGTGTATGTATACATATATATCTCT
CCATATAGGTATTTCTTTGATACTTGTAATTTTAAATTTTCAGCTTCACGAT
ATAAAATAATATAAGAACTTCTGGTTTACAAAATGTAAAATCTTAAGCCAA
TGGAACCCCTTGATTTCGATATCCAGTGTACACCCAACCTATTGGTTGTATC
AGTTTGTGTATGTGCAAATGTCAAATAATCTTTTGCTTTAATTGCTACTGT
ACTTGCTTTGAAAGATTACCTACTATTTTATGATAAAATGTAGTTGTCTCC
AGAGCTTAAATATAATTTGTAAAGCACTTGGTTTAAATTTCTCTCTACCTA
TAAACAGTTTAGCATTAAAGGGTTTCTATTAATGACACAGAATTATTGGCC
AAGTGTAATTTCTTAAAATTTAGCATTACTTTAAATAGCCAGCATGTAATA
CAAGTAACTACACTACCTCATATCTACATGATTTTCAAGTTGTAATGCAG
ATGGACAGATAAAAAAGATTTTACGTTTGTCTTTTGGCCATAAGTGGGAA
AGTTTTCTGTATATTGCATAGCATTACACATTTATGCCTATTTTAAACATTA
ACTTCTAAAGAAGTTTTTTCTAAGAAAATGTTTCAAGGCAATATTTTTTTT
GAGGCTGCCGAAGACAAATGACAGGATTATGAGTATACAGTGTATGCCT
TTTCCTTCATGCAGAATTTTGAATGTTTTTCAGTTTGTATATTGCATATTC
ACATGATCATTGCTCACTATTTTATGAACTGGCCTTCTCAATGTTTGATG
ATTTTTTAAAAGCTGTTATGTTGAATTCAGTAAAATAACATTACCTTATTT
TTTTTCTTATTCAAATTCTGGAACTATAGCAAATAATTCGTTAAATTGTCA
TATTCAAACAATGTGGATACAGTCTTGGTTCTCCATCTGTAATTTTTTTT
TAACAGTTTGCTATAGCTTACTGCTTAACTAATTTTAAATAAGGAAATAAG
TATGTTAGATGCAGTAGACGATACAGGTTGCATGTGCGCTAGCCTCGAG
TCTAGAGTCGACAGCA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

These sequencing data confirm that the directed mutation of the 1st seed region in BACH1 was successful.

7.3.1.6.2 BACH1 – 2nd Seed Mutation

FW and the antiparallel of the RV sequencing data were combined:

GTATTCTAGTTGTTTAACGGGTGTAGGGGGAGGATATTGCTAGTATATT
TTCAGTGGTTTGTATGTTCTCTCTGCTACTGACTTATTTGTAAGAGAAAA
TTAGTTGGACTTGTTTATTTTCTAGTAGCTTTTATAAGTACACTCAAGAAT
TTGTCAGGGAGAATAATTCTGATAGTGCATCCCATACTGCAAAAGAATTT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGTGTATGTATACATATATAT
CTCTCCATATAGGTATTTCTTTGATACTTGTAATTTTAAATTTTCAGCTTCA
CGATATAAAATAATATAAGAACTTCTGGTTTACAAAATGTAAAATCTTAAG
CCAATGGAACCCTTGATTTTCGATATCCAGTGTACACCCA ACTATTGGTT
GTATCAGTTTGTGTATGTGCAAATGTCAAATAATCTTTTGCTTTAATTGCT
ACTGTA CTTGCTTTGAAAGATTACCTACTATTTTATGATAAAATGTAGTTG
TCTCCAGAGCTTAAATATAATTTGTAAAGCACTTGGTTTAAATTTCTCTCT
ACCTATAAACAGTTTAGCATTAAAGGGTTTCTATTAATGACACAGAATTAT
TGGCCAAGTGAATTTCTTAAAATTTAGCATTACTTTAAATAGCCAGCAT
GTAATAAAAGTAACTACAGGATCCCATATCTACATGATTTTCAAGTTGTAA
TGCAGATGGACAGATAAAAAAGATTTTACGTTTGTCTTTTGGCCATAAGT
GGGAAAGTTTTCTGTATATTGCATAGCATTACACATTTATGCCTATTTTAA
CATTAACTTCTAAAGAAGTTTTTTCTAAGAAAATGTTTCAAGGCAATATTT
TTTTTGAGGCTGCCGAAGACAAATGACAGGATTATGAGTATACAGTGTA
TGCCTTTTCCTTCATGCAGAATTTTGAATGTTTTTCAGTTTGTATATTGCA
TATTCACATGATCATTGCTCACTATTTTATGAACTGGCCTTCTCAATGTTT
GATGATTTTTTAAAAGCTGTTATGTTGAATTCAGTAAAATAACATTACCTT
ATTTTTTTTCTTATTCAAATTCTGGA ACTATAGCAAATAATTCGTTAAATT
GTCATATTCAAAACAAATGTGGATACAGTCTTGGTTCTCCATCTGTAATT
TTTTTTAACAGTTTGCTATAGCTTACTGCTTAACTAATTTTAAATAAGGAA
ATAAGTATGTTAGATGCAGTAGACGATACAGGTTGCATGTGCGCTAGCC
TCGAGTCTAGAGTCGACTGCA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

7.3.1.6.3 CDC25Aa – Seed Mutation

FW and Antiparallel of RV sequencing combined:

CTAGTTGTTTAACGCAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCT
GTTAGACGTCCCTCCGTCCATATCAGAACTGTGCCACAATGCAGTTCTGA
GCACCGTGTCAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCCGGG
GCCTTATCGGGCTCCAGCCATCTCATGAGGGGAGAGGAGACGGAGGG
GAGTAGAGAAGTTACACAGAAATGCTGCTGGCCAAATAGCAAAGACAAC
CTGGGAAGGAAAGGTCTTTGTGGGATAATCCATATGTTTAATTTATTCAA
CTTCATCAATCACTTTATTTTATTTTTTTTTTCTAACTCCTGGAGACTTATTT
TACTGCTTCATTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTTATTAT
CCCAGGGACGGATCCGGCTTTTAATTTAAAAAAAAAAAAAGAAGTGGGTA
AGAAAATGCAAACCTGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCT
GTCACCCCCAGGAGGCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCC
AAGAACTGAGGTCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTG
ACATAAGCTAAATCTCCAGACCCACCACTGGCTACCGATATCTATTTG
GTGGGAGGTGTGGCCCTGTTCTTCCTCACCCCAGTTCCATGACATTGG
CTGGTATAGGAGCCACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTG
GGCCACCCCCGGCTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGG
GAAGGGGGGTGGGGGGTAGGTGGGCTCCACAGGATGGGGGAGGAGC
ATGTCCACTGAGTATCTTCCTTATGTTGCTGTGATATTGATAGCTTTTATT
TTCTAATTTTTAAAAATGGTCATATTATGAGTCAAAGAGTATCAAATCAG
TGTTGGATGGACCACCCAAGGGTGAGGAGAGGGGCTGGAAGCCCTGG
GCATTAGGAGAAGGGAGTGGGTGCTGGCATGGACATGACTGGATAGAA
TTTTCTCAGGAGGGAGCTTGGTGGATTTTGAAGGTAAAACCTTTCTGGGT
TTATCATGTTTTAATTTTAGAGACAGGGAGTGATGAATCATCACCGGTTG
TCCCCTTATCTAACTCCATAAAAGTGGGAATTTCAAAGAACACCTCATC
CAAGGAGCTGGGGCAGACTTCATTGATTCTAGAGAGACCTGTTTCAGTG
CCTACTCATCCCTGCCCTCTGGTGCCAGCCTCCTTACCATCACGGCTTC
ACTGAGGTGTAGGTGGGTTTTTCTTAAACAGGAGACAGTCTCTCCCCTC
TTACCTCAACTTCTTGGGGTGGGAATCAGTGATACTGGAGATGGCTAGT
TGCTGTGTTACGGGTTTGGATTACATTTGGCTATAAAACAATCTTGTTGG
GAAAAATGTGGGGGAGAGACTTCTCCTACACGCGCATTGAGACAGA
TTCCAACCTGGTTAATGATATTGTTTGTAAAGAAAGAGATTCTGTTGGTTGA

CTGCCTAAAGAGAAAGGTGGGATGGCCTTCAGATTATACCAGCTTAGCT
AGCATTACTAACCAACTGTTGGAAGCTCTGAAAATAAAAGATCTTGAACC
CATGCTCTCTGCCCGCTAGCCTCGAGTCTAGAGTCGACCGCA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

7.3.1.6.4 CDC25Ab – 1st Seed Mutation

The FW and antiparallel of RV sequences were combined:

CTAGTTGTTTAACGCAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCT
GTTAGACGTCCTCCGTCCATATCAGAACTGTGCCACAATGCAGTTCTGA
GCACCGTGTCAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCCGGG
GCCTTATCGGGCTCCAGCCATCTCATGAGGGGAGAGGAGACGGAGGG
GAGTAGAGAAGTTACACAGAAATGCTGCTGGCCAAATAGCAAAGACAAC
CTGGGAAGGAAAGGTCTTTGTGGGATAATCCATATGTTTAATTTATTCAA
CTTCATCAATCACTTTATTTTATTTTTTTTTTCTAACTCCTGGAGACTTATTT
TACTGCTTCATTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTTATTAT
CCCAGGGACGGATCCGGCTTTTAATTTAAAAAAAAAAAAAGAAGTGGGTA
AGAAAATGCAAACCTGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCT
GTCACCCCCAGGAGGCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCC
AAGAACTGAGGTCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTG
ACATAAGCTAAATCTCCAGACCCACCACTGGCTACCGATATCTATTTG
GTGGGAGGTGTGGCCCTGTTCTTCCTCACCCAGTTCCATGACATTGG
CTGGTATAGGAGCCACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTG
GGCCACCCCCGGCTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGG
GAAGGGGGGTGGGGGGTAGGTGGGCTCCACAGGATGGGGGAGGAGC
ATGTCCAAGTATCTTCCTTATGTTGCTGTGATATTGATAGCTTTTATT
TTCTAATTTTTAAAAATGGTCATATTATGAGTCAAAGAGTATCAAATCAG
TGTTGGATGGACCACCAAGGGTGAGGAGAGGGGCTGGAAGCCCTGG
GCATTAGGAGAAGGGAGTGGGTGCTGGCATGGACATGACTGGATAGAA
TTTTCTCAGGAGGGAGCTTGGTGGATTTTGAAGGTAAAACCTTCTGGGT
TTATCATGTTTTAATTTTAGAGACAGGGAGTGATGAATCATCACCGGTTG

TCCCCTTATCTAACTCCATAAAAAGTGGGAATTTCAAAGAACACCTCATC
CAAGGAGCTGGGGCAGACTTCATTGATTCTAGAGAGACCTGTTTCAGTG
CCTACTCATCCCTGCCCTCTGGTGCCAGCCTCCTTACCATCACGGCTTC
ACTGAGGTGTAGGTGGGTTTTTCTTAAACAGGAGACAGTCTCTCCCCTC
TTACCTCAACTTCTTGGGGTGGGAATCAGTGATACTGGAGATGGCTAGT
TGCTGTGTTACGGGTTTGAGTTACATTTGGCTATAAAACAATCTTGTTGG
GAAAAATGTGGGGGAGAGACTTCTTCTACACGCGCATTGAGACAGA
TTCCAAGTGGTTAATGATATTGTTTGTAAGAAAGAGATTCTGTTGGTTGA
CTGCCTAAAGAGAAAGGTGGGATGGCCTTCAGATTATACCAGCTTAGCT
AGCATTACTAACCAACTGTTGGAAGCTCTGAAAATAAAGATCTTGAACC
CATGCTCTCTGCCGCTAGCCTCGAGTCTAGAGTCGACCGCA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

These sequencing data confirm that the directed mutation of the 1st seed region of CDC25Ab was successful.

7.3.1.6.5 CDC25Ab – 2nd Seed Mutation

FW and the antiparallel of the RV sequencing data were combined:

AGTTGTTTAACGCAGGGCATCTTGCTGGCTACGCCTCTTCTGTCCCTGT
TAGACGTCCTCCGTCCATATCAGAACTGTGCCACAATGCAGTTCTGAGC
ACCGTGTCAAGCTGCTCTGAGCCACAGTGGGATGAACCAGCCGGGGC
CTTATCGGGCTCCAGCCATCTCATGAGGGGAGAGGAGACGGAGGGGA
GTAGAGAAGTTACACAGAAATGCTGCTGGCCAAATAGCAAAGACAACCT
GGGAAGGAAAGGTCTTTGTGGGATAATCCATATGTTTAATTTATTCAACT
TCATCAATCACTTTATTTTATTTTTTTTTCTAACTCCTGGAGACTTATTTA
CTGCTTCATTAGGTTGAAATACTGCCATTCTAGGTAGGGTTTTATTATCC
CAGGGACGGATCCGGCTTTTAAATTTAAAAAAGAAAGTGGGTAAG
AAAATGCAAACCTGTTATAAGTTATCGGACAGAAAGCTAGGTGCTCTGT
CACCCCAGGAGGCGCTGTGGTACTGGGGCTGCTGCTATTTAAGCCAA
GAACTGAGGTCCTGGTGAGAGCGTTGGACCCAGGCTTGGCTGCCTGAC

ATAAGCTAAATCTCCCAGACCCACCACTGGCTACCGATATCTATTTGGT
 GGGAGGTGTGGCCCTGTTCTTCCTCACCCCAGTTCCATGACATTGGCT
 GGTATAGGAGCCACAGTCAGGAAAGCACTTGAGGCAGCATCTGTTGGG
 CCACCCCGGCTCAGTGCTGGAATGTTGCAGTGTAGGTTTCCCAGGGA
 AGGGGGGTGGGGGTAGGTGGGCTCCACAGGATGGGGGAGGAGCATG
 TCCACTGAGTATCTTCCTTATGTTGCTGTGATATTGATAGCTTTTATTTTC
 TAATTTTTAAAAATGGTCATATTATGAGTCAAAGAGTATCAAATCAGTGT
 TGGATGGACCACCCAAGGGTGAGGAGAGGGGCTGGAAGCCCTGGGCA
 TTAGGAGAAGGGAGTGGGTGCTGGCATGGACATGACTGGATAGAATTT
 TCTCAGGAGGGAGCTTGGTGGATTTTGAAGGTAAAACCTTCTGGGTTTA
 TCATGTTTTAATTTTAGAGACAGGGAGTGATGAATCATCACCGGTTGTC
 CCCTTATCTAACTCCATAAAAGTGGGAATTTCAAAGAACACCTCATCCA
 AGGAGCTGGGGCAGACTTCATTGATTCTAGAGAGACCTGTTTCAGTGC
 CTACTCATCCCTGCCCTCTGGTGCCAGCCTCCTTACCATCACGGCTTCA
 CTGAGGTGTAGGTGGGTTTTTCTTAAACAGGAGACAGTCTCTCCCCTCT
 CAATTGAACTTCTTGGGGTGGGAATCAGTGATACTGGAGATGGCTAGTT
 GCTGTGTTACGGGTTTGAGTTACATTTGGCTATAAAACAATCTTGTTGG
 GAAAAATGTGGGGGAGAGGACTTCTTCCTACACGCGCATTGAGACAGA
 TTCCAACCTGGTTAATGATATTGTTTGTAAGAAAGAGATTCTGTTGGTTGA
 CTGCCTAAAGAGAAAGGTGGGATGGCCTTCAGATTATAACCAGCTTAGCT
 AGCATTACTAACCAACTGTTGGAAGCTCTGAAAATAAAAGATCT**TTGAACC**
CATGCTCTCTGCCCGCTAGCCTCGAGTCTAGAGTCGACCTGCAG

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

7.3.1.6.6 *HMGA2b – 1st Seed Mutation*

FW sequencing only:

TTGTTTAACG**GCTGCTTCAGGGAGGTAGTTTCAAAGGCCACA****GATATCT**
 CTGAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTATG

GAGAGAATTA AAAACTCAAC TTA CTGTTAACTGTGCGTTAAATAAGCAA
TAAACAGTGGCTCATAAAAATAAAAAGTCGCATTCCATATCTTTGGATGGG
CCTTTTAGAAACCTCATTGGCCAGCTCATAAAATGGAAGCAATTGCTCAT
GTTGGCCAAACATGGTGCACCGAGTGATTTCCATCTCTGGTAAAGTTAC
ACTTTTATTTCTGTATGTTGTACAATCAAACACACTACTACCTCTTAAG
TCCCAGTATACCTCATT TTTTCATACTGAAAAAAAAAAGCTTGTGGCCAATG
GAACAGTAAGAACATCATAAAATTTTATATATATAGTTTATTTTTGTGGG
AGATAAATTTTATAGGACTG **TTCTTTGCTGTTGTTGGTCGC**CGCTAGCCT
CGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAAC
AAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAA
CTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATGATAAGATA
CATTGATGAGTTTGGACAAACCACA ACTAGAATGCAGTGAAAAAATGC
TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGC
TGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTT
CAGGGGGAGATGTGGGAGGTTTTTTAAGCAAGTAAACCTCTACAAAT
GTGGTAAAATCGAATTTTAAACAAAATATTAACGCTTACAATTTCTCTGATG
CGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGCG
GATCTGCGCAGCACCATGGCCTGAAATAACCTCTGAAAGAGG

- Start of insert
- End of insert
- **Seed region**
- **Successfully mutated site**

These sequencing data confirm that the directed mutation of the 1st seed region of HMGA2b was successful.

7.3.1.6.7 *HMGA2b – 2nd Seed Mutation*

FW sequencing only:

TATCTAGTTGTTTAACGGCTGCTTCAGGGAGGTAGTTTCAAAGGCCACA
GATATCTCTGAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGA

GCTATGGAGAGAATTAAACTCAACATTACTGTAACTGTGCGTTAAATA
 AGCAAATAAACAGTGGCTCATAAAAATAAAAGTCGCATTCCATATCTTTG
 GATGGGCCTTTTAGAAACCTCATTGGCCAGCTCATAAAATGGAAGCAAT
 TGCTCATGTTGGCCAAACATGGTGCACCGAGTGATTTCCATCTCTGGTA
 AAGTTACACTTTTATTTCTGTATGTTGTACAAT**CAAAACACACTACGGA**
TCCTTAAGTCCCAGTATACCTCATTTTTCATACTGAAAAAAAAAGCTTGT
 GGCCAATGGAACAGTAAGAACATCATAAAATTTTTATATATATAGTTTATT
 TTTGTGGGAGATAAATTTTATAGGACTG**TTCTTTGCTGTTGTTGGTCGCC**
 GCTAGCCTCGAGTCTAGAGTCGACCTGCAGGCATGCAAGCTGATCCGG
 CTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTG
 AGCAATAACTAGCATAACCCCTTGGGGCGGCCGCTTCGAGCAGACATG
 ATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAA
 AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA
 TTATAAGCTGCAATAACAAGTTAACAACAACAATTGCATTCATTTTATGT
 TTCAGGTTTCAGGGGGAGATGTGGGAGGTTTTTTTTAAGCAAGTAAACCT
 CTACAAATGTGGTAAAATCGAATTTTAACAAAATATTAACGCTTACAATTT
 CCTGATGCGGTATTTTCTCCTTACGCATCTGTGC

- Start of insert
- End of insert
- **Seed region**
- **Successfully mutated site**

7.3.1.6.8 *HMGA2b – 3rd Seed Mutation*

Antiparallel of RV sequencing only:

CCCATGATCATGAGCGGCTACGTTAACACCCCGAGGCTACAAACGCTC
 TCATCGACAAGGACGGCTGGCTGCACAGCGGCGACATCGCCTACTGG
 GACGAGGACGAGCACTTCTTCATCGTGGACCGGCTGAAGAGCCTGATC
 AAATACAAGGGCTACAGGTAGCCCCAGCCGAACTGGAGAGCATCGTGC
 TGCAACACCCCAACATCTTCGACGCCGGGGTCCGCCGGCCTGCCCGAC
 GACGATGCCGGCGAGCTGCCCGCCGACGTCGTCGTGCTGGAACACGG
 TAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGT
 TACAACCGCCAAGAAGCTGCGCGGTGGTGTGTTGTTTCGTGGACGAGGT
 GCCTAAAGGACTGACCGGCAAGTTGGACGCCCGCAAGATCCGCGAGAT

TTCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAATTCTAGTTG
TTAAACGGCTGCTTCAGGGAGGTAGTTCAAAGGCCACAGATATCTCT
GAGACTGGCAGATCGCTCACTGTTGTGAATCACCAAAGGAGCTATGGA
GAGAATTAAACTCAACATTACTGTAACTGTGCGTTAAATAAGCAAATA
AACAGTGGCTCATAAAAATAAAAGTCGCATTCCATATCTTTGGATGGGC
CTTTTAGAAACCTCATTGGCCAGCTCATAAAATGGAAGCAATTGCTCAT
GTTGGCCAAACATGGTGCACCGAGTGATTTCCATCTCTGGTAAAGTTAC
ACTTTTATTTCTGTATGTTGTACAATCAAACACACTACGGATCCTTAA
GTCCCAGTAAGATCTATTTTTTCACTGAAAAAAAAAAGCTTGTGGCCAAT
GGAACAGTAAGAACATCATAAAATTTTTATATATATAGTTTATTTTTGTGG
GAGATAAATTTATAGGACTGTTCTTTGCTGTTGTTGGTCGCCGCTAGC
CTCGAGTCTAGAGTCGACCTGCA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

7.3.1.6.9 KRAS – Seed Mutation

Antiparallel of RV sequencing only:

CGAGATTCTCATTAAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAATT
CTAGTTGTTTAAACGGTCTCTGTGCCAGCTCTATAATTGTTTTGCTACG
ATTCCACTGAACTCTTCGATCAAGCTACTTTATGTAAATCACTTCATTG
TTTTAAAGGAATAAACTTGATTATATTGTTTTTTATTTGGCATAACTGTG
ATTCTTTTGGGACAATACTGTACACATTAAGGTGTATGTCAGATATTCA
TATTGACCCAAATGTGTAATATTCCAGTTTTCTCTGCATAAGTAATTA
TATACTTAAAAATTAATAGTTTTATCTGGGTACAAATAAACAGGTGCCTG
AACTAGTTCACAGACAAGGAACTTCTATGTAAAAATCACTATGATTTCT
GAATTGCTATGTGAACTACAGATCTTTGGAACACTGTTTAGGTAGGGT
GTTAAGACTTACACAGGATATC GATATC GTTTCTACACAGAGAAAGAAATGGCCA
TACTTCAGGAAGTGCAGTGCTTATGAGGGGATATTTAGGCCTCTTGAAT
TTTTGATGTAGATGGGCATTTTTTTAAGGTAGTGGTAATTACCTTTATGT
GAACTTTGAATGGTTTAAACAAAAGATTTGTTTTTGTAGAGATTTTAAAGG
GGGAGAATTCTAGAAATAAATGTTACCTAATTATTACAGCCTTAAAGATA

AAAATCCTTGTTGAAGTTTTTTAAAAAAGCTAAATTACATAGACTTAGG
CATTACATGTTTGTGGAAGAATATAGCAGACGTATATTGTATCATTG
GTGAATGTTCCAAGTAGGCACGCTAGCCTCGAGTCTAGAGTCGACAT
C

- Start of insert
- End of insert
- **Seed region**
- **Successfully mutated site**

These sequencing data confirm that the directed mutation of the KRAS seed region was successful.

7.3.1.6.10 MYC – Seed Mutation

FW sequencing only:

CTAGTTGTTTAACGA**AATGTCCTGAGCAATCAC**CTATGAACTTGTTTCAA
TGCATGATCAAATGCAACCTCACAACCTTGGCTGAGTCTTGAGACTGAA
AGATTTA**GCCATAATGTAAGATATCTC**CGCTAGCCTCGAGTCTAGAGTC
GACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAAAGCCCGAAAGG
AAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCC
TTGGGGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTT
GGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAAT
TTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAG
TTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGGAGATG
TGGGAGGTTTTTTAAGCAAGTAAACCTCTACAAATGTGGTAAAATCGA
ATTTTAACAAAATATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCT
TACGCATCTGTGCGGTATTTACACCCGCATACGCGGATCTGCGCAGCA
CCATGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCT
GAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAA
AGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCA
ATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCA
GAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCC
CCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTC
TCCGCCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCC

GCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGA
 GGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGA

- Start of insert
- End of insert
- Seed region
- Successfully mutated site

7.3.2 Luciferase Assay Optimisation

Caco-2 cells were seeded in 96-well plates with 20 000 cells/well, and simultaneously 'fast-forward' transfected with the 'CDC25Ab 5 WT' vector using the Attractene Transfection Reagent, with varying amounts of vector and Attractene to determine the optimum combination.

The luciferase assay was done 2 days after seeding/transfection:

7.3.2.1 Firefly Data (Blank-Corrected)

DNA (µg)	None	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3
Attr. (µL)	None	0.25	0.375	0.75	0.5	0.75	1.5	0.75	1.125	2.25
		15973	12708	19347	18360	24862	7437	18813	25022	5856
		16995	13110	18944	21603	25792	4996	9317	23554	2638
		16561	15563	19330	22239	23115	7730	21013	23523	4428
Mean		16510	13794	19207	20734	24590	6721	16381	24033	4308
S.E.M.		242	729	107	981	641	708	2930	404	760

7.3.2.2 Renilla Data (Blank-Corrected)

DNA (µg)	None	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3
Attr. (µL)	None	0.25	0.375	0.75	0.5	0.75	1.5	0.75	1.125	2.25
		15973	12708	19347	18360	24862	7437	18813	25022	5856
		16995	13110	18944	21603	25792	4996	9317	23554	2638
		16561	15563	19330	22239	23115	7730	21013	23523	4428
Mean		16510	13794	19207	20734	24590	6721	16381	24033	4308
S.E.M.		242	729	107	981	641	708	2930	404	760

7.3.2.3 Firefly/Renilla

DNA (µg)	None	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3
Attr. (µL)	None	0.25	0.375	0.75	0.5	0.75	1.5	0.75	1.125	2.25
		0.90	0.97	0.94	1.04	0.95	1.08	1.02	0.97	0.97
		1.01	1.06	0.98	1.01	0.94	1.07	1.41	0.97	1.23
		0.96	0.98	1.07	1.02	1.04	1.07	1.13	1.05	0.91
	Mean	0.96	1.01	1.00	1.02	0.97	1.07	1.19	0.99	1.04
	S.E.M.	0.03	0.02	0.03	0.01	0.03	0.00	0.09	0.02	0.08

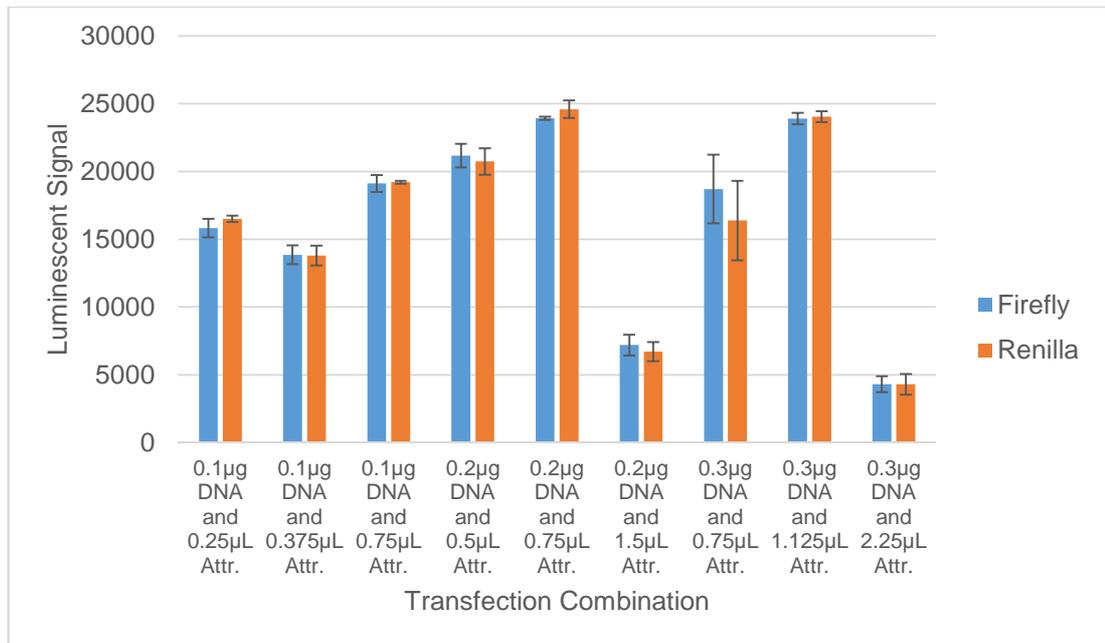


Figure 114: The luminescent signals from the experiments to optimise the amounts of DNA and Attractene transfection reagent to use are represented in this graph as means \pm S.E.M.

It is clear from the data illustrated in Figure 114 that the optimal amounts of DNA and Attractene to use per well are 0.2µg and 0.75µL respectively, since these give the highest luminescent signals, hence these will be used for experiments. It is also clear that using more than 1.125µL Attractene/well results in significant cytotoxicity.

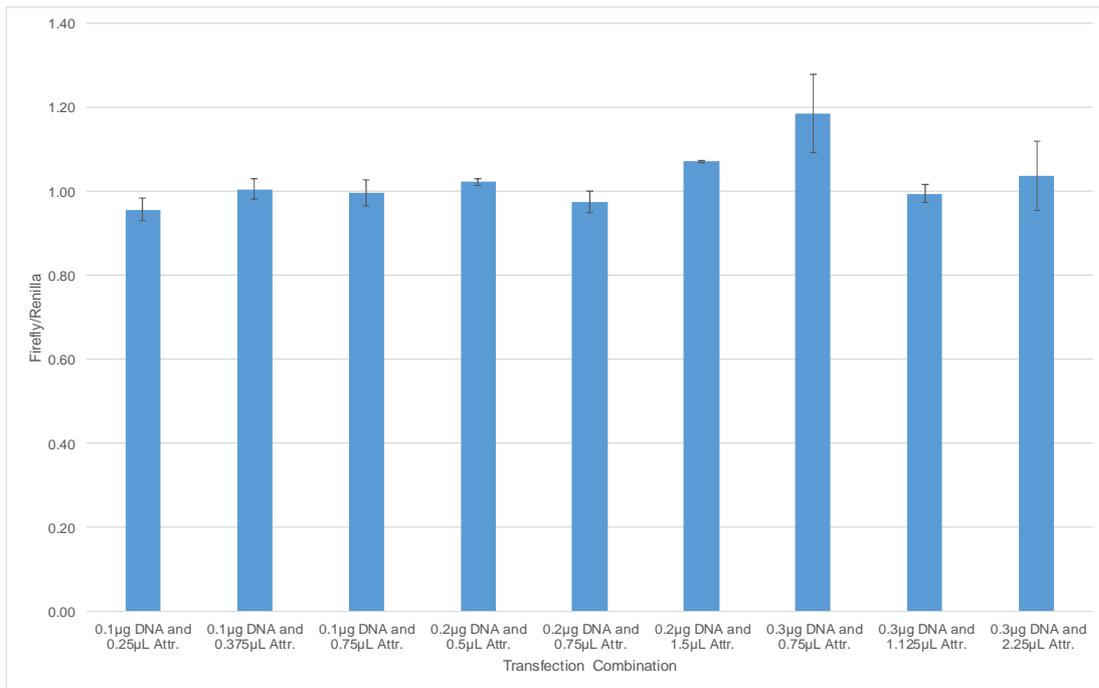


Figure 115: The [Firefly luminescence]/[Renilla luminescence] ratios are represented as means ± S.E.M.

The data in Figure 115 confirm that as expected, varying the amounts of DNA and Attractene did not affect Firefly/*Renilla* ratios, indicating that the measurement of *Renilla* luciferase activity is suitable for use as an internal control.

Bibliography

Adenuga, D., S. Caito, H. Yao, I. K. Sundar, J.-W. Hwang, S. Chung and I. Rahman (2010). "Nrf2 deficiency influences susceptibility to steroid resistance via HDAC2 reduction." Biochemical and Biophysical Research Communications **403**(3–4): 452-456.

Akhtar, N., Z. Rasheed, S. Ramamurthy, A. N. Anbazhagan, F. R. Voss and T. M. Haqqi (2010). "MicroRNA-27b regulates the expression of matrix metalloproteinase 13 in human osteoarthritis chondrocytes." Arthritis & Rheumatism **62**(5): 1361-1371.

Amirkhah, R., U. Schmitz, M. Linnebacher, O. Wolkenhauer and A. Farazmand (2015). "MicroRNA-mRNA interactions in colorectal cancer and their role in tumor progression." Genes Chromosomes Cancer **54**(3): 129-141.

Andréasson, E., L. B. Jørgensen, A. Höglund, L. Rask and J. Meijer (2001). "Different Myrosinase and Idioblast Distribution in Arabidopsis and *Brassica napus*." Plant Physiology **127**(4): 1750-1763.

Arlt, A., I. Bauer, C. Schafmayer, J. Tepel, S. S. Muerkoster, M. Brosch, C. Roder, H. Kalthoff, J. Hampe, M. P. Moyer, U. R. Folsch and H. Schafer (2009). "Increased proteasome subunit protein expression and proteasome activity in colon cancer relate to an enhanced activation of nuclear factor E2-related factor 2 (Nrf2)." Oncogene **28**(45): 3983-3996.

Bäckdahl, L., A. Bushell and S. Beck (2009). "Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation." The International Journal of Biochemistry & Cell Biology **41**(1): 176-184.

Bai, Y., X. Wang, S. Zhao, C. Ma, J. Cui and Y. Zheng (2015). "Sulforaphane Protects against Cardiovascular Disease via Nrf2 Activation." Oxidative Medicine and Cellular Longevity **2015**: 13.

Baraniskin, A., K. Birkenkamp-Demtroder, A. Maghnouj, H. Zöllner, J. Munding, S. Klein-Scory, A. Reinacher-Schick, I. Schwarte-Waldhoff, W. Schmiegel and S. A. Hahn (2012). "MiR-30a-5p suppresses tumor growth in colon carcinoma by targeting DTL." Carcinogenesis **33**(4): 732-739.

Bartel, D. P. (2004). "MicroRNAs: Genomics, Biogenesis, Mechanism, and Function." Cell **116**(2): 281-297.

Basu, A., H. Alder, A. Khiyami, P. Leahy, C. M. Croce and S. Haldar (2011). "MicroRNA-375 and MicroRNA-221: Potential Noncoding RNAs Associated with Antiproliferative Activity of Benzyl Isothiocyanate in Pancreatic Cancer." Genes & Cancer **2**(2): 108-119.

Bat-Chen, W., T. Golan, I. Peri, Z. Ludmer and B. Schwartz (2010). "Allicin Purified From Fresh Garlic Cloves Induces Apoptosis in Colon Cancer Cells Via Nrf2." Nutrition and Cancer **62**(7): 947-957.

Bonfrate, L., D. F. Altomare, M. Di Lena, E. Travaglio, M. T. Rotelli, A. De Luca and P. Portincasa (2013). "MicroRNA in Colorectal Cancer: New Perspectives for Diagnosis, Prognosis and Treatment." Journal of Gastrointestinal and Liver Diseases **22**(3): 311-320.

Boutros, R., V. Lobjois and B. Ducommun (2007). "CDC25 phosphatases in cancer cells: key players? Good targets?" Nat Rev Cancer **7**(7): 495-507.

Burk, U., J. Schubert, U. Wellner, O. Schmalhofer, E. Vincan, S. Spaderna and T. Brabletz (2008). A reciprocal repression between ZEB1 and members of the miR - 200 family promotes EMT and invasion in cancer cells.

Burow, M., A. Bergner, J. Gershenzon and U. Wittstock (2007). "Glucosinolate hydrolysis in *Lepidium sativum*—identification of the thiocyanate-forming protein." Plant Molecular Biology **63**(1): 49-61.

Calin, G. A. and C. M. Croce (2006). "MicroRNA-Cancer Connection: The Beginning of a New Tale." Cancer Research **66**(15): 7390-7394.

Chang, C.-J., D. J. Mulholland, B. Valamehr, S. Mosessian, W. R. Sellers and H. Wu (2008). "PTEN Nuclear Localization Is Regulated by Oxidative Stress and Mediates p53-Dependent Tumor Suppression." Molecular and Cellular Biology **28**(10): 3281-3289.

Chen, B., Y. Liu, X. Jin, W. Lu, J. Liu, Z. Xia, Q. Yuan, X. Zhao, N. Xu and S. Liang (2014). "MicroRNA-26a regulates glucose metabolism by direct targeting PDHX in colorectal cancer cells." BMC Cancer **14**: 443.

Chen, C.-Y. A. and A.-B. Shyu (1995). "AU-rich elements: characterization and importance in mRNA degradation." Trends in Biochemical Sciences **20**(11): 465-470.

Chen, M. J., W. Y. Tang, C. W. Hsu, Y. T. Tsai, J. F. Wu, C. W. Lin, Y. M. Cheng and Y. C. Hsu (2012). "Apoptosis Induction in Primary Human Colorectal Cancer Cell Lines and Retarded Tumor Growth in SCID Mice by Sulforaphane." Evid Based Complement Alternat Med **2012**: 415231.

Chen, X.-L., G. Dodd, S. Thomas, X. Zhang, M. A. Wasserman, B. H. Rovin and C. Kunsch (2006). "Activation of Nrf2/ARE pathway protects endothelial cells from oxidant injury and inhibits inflammatory gene expression." American Journal of Physiology - Heart and Circulatory Physiology **290**(5): H1862-H1870.

Cheng, X., C.-H. Ku and R. C. M. Siow (2013). "Regulation of the Nrf2 antioxidant pathway by microRNAs: New players in micromanaging redox homeostasis." Free Radical Biology and Medicine **64**(0): 4-11.

Chiyomaru, T., H. Enokida, S. Tatarano, K. Kawahara, Y. Uchida, K. Nishiyama, L. Fujimura, N. Kikkawa, N. Seki and M. Nakagawa (2010). "miR-145 and miR-133a function as tumour suppressors and directly regulate FSCN1 expression in bladder cancer." British journal of cancer **102**(5): 883-891.

Cho, W., J. Shin, J. Kim, M. Lee, K. Hong, J.-H. Lee, K. Koo, J. Park and K.-S. Kim (2009). "miR-372 regulates cell cycle and apoptosis of ags human gastric cancer cell line through direct regulation of LATS2." Molecules and Cells **28**(6): 521-527.

Chua, H. L., P. Bhat-Nakshatri, S. E. Clare, A. Morimiya, S. Badve and H. Nakshatri (2006). "NF-[kappa]B represses E-cadherin expression and enhances epithelial to mesenchymal transition of mammary epithelial cells: potential involvement of ZEB-1 and ZEB-2." Oncogene **26**(5): 711-724.

Cifuentes, D., H. Xue, D. W. Taylor, H. Patnode, Y. Mishima, S. Cheloufi, E. Ma, S. Mane, G. J. Hannon, N. D. Lawson, S. A. Wolfe and A. J. Giraldez (2010). "A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity." Science **328**(5986): 1694-1698.

Cittelly, D. M., P. M. Das, N. S. Spoelstra, S. M. Edgerton, J. K. Richer, A. D. Thor and F. E. Jones (2010). "Downregulation of miR-342 is associated with tamoxifen resistant breast tumors." Molecular cancer **9**: 317.

Clarke, J. D., R. H. Dashwood and E. Ho (2008). "Multi-targeted prevention of cancer by sulforaphane." Cancer Lett **269**(2): 291-304.

Cloonan, N., S. Wani, Q. Xu, J. Gu, K. Lea, S. Heater, C. Barbacioru, A. L. Steptoe, H. C. Martin, E. Nourbakhsh, K. Krishnan, B. Gardiner, X. Wang, K. Nones, J. A. Steen, N. A. Matigian, D. L. Wood, K. S. Kassahn, N. Waddell, J. Shepherd, C. Lee, J. Ichikawa, K. McKernan, K. Bramlett, S. Kuersten and S. M. Grimmond (2011). "MicroRNAs and their isomiRs function cooperatively to target common biological pathways." Genome Biol **12**(12): R126.

Cullinan, S. B., J. D. Gordan, J. Jin, J. W. Harper and J. A. Diehl (2004). "The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase." Mol Cell Biol **24**(19): 8477-8486.

Cunningham, D., W. Atkin, H. Lenz, H. T. Lynch, B. Minsky, B. Nordlinger and N. Starling (2010). "Colorectal cancer." The Lancet **375**(9719): 1030-1047.

Davidson, R. K., O. Jupp, Y. Bao, A. J. MacGregor, S. T. Donell, A. Cassidy and I. M. Clark (2016). "Can sulforaphane prevent the onset or slow the progression of osteoarthritis?" Nutrition Bulletin **41**(2): 175-179.

Derks, S., L. J. Bosch, H. E. Niessen, P. T. Moerkerk, S. M. van den Bosch, B. Carvalho, S. Mongera, J. W. Voncken, G. A. Meijer, A. P. de Bruine, J. G. Herman and M. van Engeland (2009). "Promoter CpG island hypermethylation- and H3K9me3 and H3K27me3-mediated epigenetic silencing targets the deleted in colon cancer (DCC) gene in colorectal carcinogenesis without affecting neighboring genes on chromosomal region 18q21." Carcinogenesis **30**(6): 1041-1048.

DIANA-Lab. "microT-CDS." Retrieved 27 Feb 14, 2014, from http://diana.cslab.ece.ntua.gr/micro-CDS/index.php?r=search/results_mature&mir=hsa-miR-27b-5p&kwd=MIMAT0004588.

Duan, Z., E. Choy, D. Harmon, C. Yang, K. Ryu, J. Schwab, H. Mankin and F. J. Hornicek (2009). "ZNF93 increases resistance to ET-743 (Trabectedin; Yondelis) and PM00104 (Zalypsis) in human cancer cell lines." PLoS One **4**(9): e6967.

Fearon, E. R. and B. Vogelstein (1990). "A genetic model for colorectal tumorigenesis." Cell **61**(5): 759-767.

Ferlay, J., H. R. Shin, F. Bray, D. Forman, C. Mathers and D. M. Parkin. (2010). "Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 " Retrieved 23rd August, 2013, from <http://globocan.iarc.fr>.

Filippova, M., V. Filippov, V. M. Williams, K. Zhang, A. Kokoza, S. Bashkirova and P. Duerksen-Hughes (2014). "Cellular levels of oxidative stress affect the response of cervical cancer cells to chemotherapeutic agents." Biomed Res Int **2014**: 574659.

Frank, S. A. (2007). Dynamics of Cancer: Incidence, Inheritance, and Evolution. Princeton (NJ).

Franklin, J. L., C. R. Rankin, S. Levy, J. R. Snoddy, B. Zhang, M. K. Washington, J. M. Thomson, R. H. Whitehead and R. J. Coffey (2013). "Malignant transformation of colonic epithelial cells by a colon-derived long noncoding RNA." Biochemical and Biophysical Research Communications **440**(1): 99-104.

Fukunaga, R., Bo W. Han, J.-H. Hung, J. Xu, Z. Weng and Phillip D. Zamore (2012). "Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals." Cell **151**(3): 533-546.

- Gan, N., X. Sun and L. Song (2010). "Activation of Nrf2 by Microcystin-LR Provides Advantages for Liver Cancer Cell Growth." Chemical Research in Toxicology **23**(9): 1477-1484.
- Gantier, M. P., C. E. McCoy, I. Rusinova, D. Saulep, D. Wang, D. Xu, A. T. Irving, M. A. Behlke, P. J. Hertzog, F. Mackay and B. R. G. Williams (2011). "Analysis of microRNA turnover in mammalian cells following Dicer1 ablation." Nucleic Acids Research **39**(13): 5692-5703.
- Gonzalez, C. D., S. Alvarez, A. Ropolo, C. Rosenzvit, M. F. Bagnes and M. I. Vaccaro (2014). "Autophagy, Warburg, and Warburg reverse effects in human cancer." Biomed Res Int **2014**: 926729.
- Gopalakrishna, R. and U. Gundimeda (2002). "Antioxidant regulation of protein kinase C in cancer prevention." J Nutr **132**(12): 3819S-3823S.
- Gregory, P. A., A. G. Bert, E. L. Paterson, S. C. Barry, A. Tsykin, G. Farshid, M. A. Vadas, Y. Khew-Goodall and G. J. Goodall (2008). "The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1." Nat Cell Biol **10**(5): 593-601.
- Guo, L., R. Yang, Z. Wang, Q. Guo and Z. Gu (2014). "Glucoraphanin, sulforaphane and myrosinase activity in germinating broccoli sprouts as affected by growth temperature and plant organs." Journal of Functional Foods **9**: 70-77.
- Haggar, F. A. and R. P. Boushey (2009). "Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors." Clin Colon Rectal Surg **22**(4): 191-197.
- Hammell, C. M. (2008). "The microRNA-argonaute complex: A platform for mRNA modulation." RNA Biology **5**(3): 123-127.
- Harman, D. (1956). "Aging: A Theory Based on Free Radical and Radiation Chemistry." Journal of Gerontology **11**(3): 298-300.
- Hata, A. and J. Lieberman (2015). "Dysregulation of microRNA biogenesis and gene silencing in cancer." Sci Signal **8**(368): re3.
- Helwak, A., G. Kudla, T. Dudnakova and D. Tollervey (2013). "Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding." Cell **153**(3): 654-665.
- Higgins, L. G., M. O. Kelleher, I. M. Eggleston, K. Itoh, M. Yamamoto and J. D. Hayes (2009). "Transcription factor Nrf2 mediates an adaptive response to sulforaphane that protects fibroblasts in vitro against the cytotoxic effects of electrophiles, peroxides and redox-cycling agents." Toxicology and Applied Pharmacology **237**(3): 267-280.

Hildebrandt, M. A. T., J. Gu, J. Lin, Y. Ye, W. Tan, P. Tamboli, C. G. Wood and X. Wu (2010). "Hsa-miR-9 methylation status is associated with cancer development and metastatic recurrence in patients with clear cell renal cell carcinoma." Oncogene **29**(42): 5724-5728.

Hiltunen, M. O., L. Alhonen, J. Koistinaho, S. Myohanen, M. Paakkonen, S. Marin, V. M. Kosma and J. Janne (1997). "Hypermethylation of the APC (adenomatous polyposis coli) gene promoter region in human colorectal carcinoma." Int J Cancer **70**(6): 644-648.

Hintze, K. J., Y. Katoh, K. Igarashi and E. C. Theil (2007). "Bach1 Repression of Ferritin and Thioredoxin Reductase1 Is Heme-sensitive in Cells and in Vitro and Coordinates Expression with Heme Oxygenase1, β -Globin, and NADP(H) Quinone (Oxido) Reductase1." Journal of Biological Chemistry **282**(47): 34365-34371.

Hitchler, M. J. and F. E. Domann (2007). "An epigenetic perspective on the free radical theory of development." Free Radical Biology and Medicine **43**(7): 1023-1036.

Homma, S., Y. Ishii, Y. Morishima, T. Yamadori, Y. Matsuno, N. Haraguchi, N. Kikuchi, H. Satoh, T. Sakamoto, N. Hizawa, K. Itoh and M. Yamamoto (2009). "Nrf2 enhances cell proliferation and resistance to anticancer drugs in human lung cancer." Clin Cancer Res **15**(10): 3423-3432.

Hou, W., Q. Tian, N. M. Steuerwald, L. W. Schrum and H. L. Bonkovsky (2012). "The let-7 microRNA enhances heme oxygenase-1 by suppressing Bach1 and attenuates oxidant injury in human hepatocytes." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1819**(11-12): 1113-1122.

Hsu, A., C. P. Wong, Z. Yu, D. E. Williams, R. H. Dashwood and E. Ho (2011). "Promoter de-methylation of cyclin D2 by sulforaphane in prostate cancer cells." Clin Epigenetics **3**(3).

Hu, H., X. Zhao, Z. Jin and M. Hou (2015). "Hsa-let-7g miRNA regulates the anti-tumor effects of gastric cancer cells under oxidative stress through the expression of DDR genes." The Journal of Toxicological Sciences **40**(3): 329-338.

Huxley, R. R., A. Ansary-Moghaddam, P. Clifton, S. Czernichow, C. L. Parr and M. Woodward (2009). "The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence." Int J Cancer **125**(1): 171-180.

Ito, K., T. Hanazawa, K. Tomita, P. J. Barnes and I. M. Adcock (2004). "Oxidative stress reduces histone deacetylase 2 activity and enhances IL-8 gene expression: role of tyrosine nitration." Biochemical and Biophysical Research Communications **315**(1): 240-245.

Izzotti, A., P. Larghero, C. Cartiglia, M. Longobardi, U. Pfeffer, V. E. Steele and S. De Flora (2010). "Modulation of microRNA expression by budesonide, phenethyl isothiocyanate and cigarette smoke in mouse liver and lung." Carcinogenesis **31**(5): 894-901.

Janobi, A. A. A., R. F. Mithen, A. V. Gasper, P. N. Shaw, R. J. Middleton, C. A. Ortori and D. A. Barrett (2006). "Quantitative measurement of sulforaphane, iberin and their mercapturic acid pathway metabolites in human plasma and urine using liquid chromatography–tandem electrospray ionisation mass spectrometry." Journal of Chromatography B **844**(2): 223-234.

Jiang, S., H.-W. Zhang, M.-H. Lu, X.-H. He, Y. Li, H. Gu, M.-F. Liu and E.-D. Wang (2010). "MicroRNA-155 Functions as an OncomiR in Breast Cancer by Targeting the Suppressor of Cytokine Signaling 1 Gene." Cancer Research **70**(8): 3119-3127.

Jin, H., W. Gong, C. Zhang and S. Wang (2013). "Epigallocatechin gallate inhibits the proliferation of colorectal cancer cells by regulating Notch signaling." Onco Targets Ther **6**: 145-153.

Kallifatidis, G., S. Labsch, V. Rausch, J. Mattern, J. Gladkich, G. Moldenhauer, M. W. Buchler, A. V. Salnikov and I. Herr (2011). "Sulforaphane increases drug-mediated cytotoxicity toward cancer stem-like cells of pancreas and prostate." Mol Ther **19**(1): 188-195.

Keum, Y. S., S. Yu, P. P. Chang, X. Yuan, J. H. Kim, C. Xu, J. Han, A. Agarwal and A. N. Kong (2006). "Mechanism of action of sulforaphane: inhibition of p38 mitogen-activated protein kinase isoforms contributing to the induction of antioxidant response element-mediated heme oxygenase-1 in human hepatoma HepG2 cells." Cancer Res **66**(17): 8804-8813.

Kim, B.-R., R. Hu, Y.-S. Keum, V. Hebbar, G. Shen, S. S. Nair and A.-N. T. Kong (2003). "Effects of Glutathione on Antioxidant Response Element-Mediated Gene Expression and Apoptosis Elicited by Sulforaphane." Cancer Research **63**(21): 7520-7525.

Kim, Y. J., S. H. Hwang, S. Y. Lee, K. K. Shin, H. H. Cho, Y. C. Bae and J. S. Jung (2012). "miR-486-5p induces replicative senescence of human adipose tissue-derived mesenchymal stem cells and its expression is controlled by high glucose." Stem cells and development **21**(10): 1749-1760.

Kohler, U. A., S. Kurinna, D. Schwitter, A. Marti, M. Schafer, C. Hellerbrand, T. Speicher and S. Werner (2014). "Activated Nrf2 impairs liver regeneration in mice by activation of genes involved in cell-cycle control and apoptosis." Hepatology **60**(2): 670-678.

Kombairaju, P., J. Ma, R. K. Thimmulappa, S. G. Yan, E. Gabrielson, A. Singh and S. Biswal (2012). "Prolonged sulforaphane treatment does not

enhance tumorigenesis in oncogenic K-ras and xenograft mouse models of lung cancer." J Carcinog **11**: 8.

Koroleva, O. A., A. Davies, R. Deeken, M. R. Thorpe, A. D. Tomos and R. Hedrich (2000). "Identification of a New Glucosinolate-Rich Cell Type in Arabidopsis Flower Stalk." Plant Physiology **124**(2): 599-608.

Kourtidis, A., S. P. Ngok, P. Pulimeno, R. W. Feathers, L. R. Carpio, T. R. Baker, J. M. Carr, I. K. Yan, S. Borges, E. A. Perez, P. Storz, J. A. Copland, T. Patel, E. A. Thompson, S. Citi and P. Z. Anastasiadis (2015). "Distinct E-cadherin-based complexes regulate cell behaviour through miRNA processing or Src and p120 catenin activity." Nat Cell Biol **17**(9): 1145-1157.

Kurinna, S. and S. Werner (2015). "NRF2 and microRNAs: new but awaited relations." Biochemical Society Transactions **43**(5): 595-601.

La Marca, M., P. Befly, C. Della Croce, P. G. Gervasi, R. Iori, E. Puccinelli and V. Longo (2012). "Structural influence of isothiocyanates on expression of cytochrome P450, phase II enzymes, and activation of Nrf2 in primary rat hepatocytes." Food and Chemical Toxicology **50**(8): 2822-2830.

Lam, T. K., L. Gallicchio, K. Lindsley, M. Shiels, E. Hammond, X. G. Tao, L. Chen, K. A. Robinson, L. E. Caulfield, J. G. Herman, E. Guallar and A. J. Alberg (2009). "Cruciferous vegetable consumption and lung cancer risk: a systematic review." Cancer Epidemiol Biomarkers Prev **18**(1): 184-195.

Lao, V. V. and W. M. Grady (2011). "Epigenetics and colorectal cancer." Nature reviews. Gastroenterology & hepatology **8**(12): 686-700.

Lau, A., N. F. Villeneuve, Z. Sun, P. K. Wong and D. D. Zhang (2008). "Dual roles of Nrf2 in cancer." Pharmacol Res **58**(5-6): 262-270.

Lee, J. Y., H. J. Kim, N. A. Yoon, W. H. Lee, Y. J. Min, B. K. Ko, B. J. Lee, A. Lee, H. J. Cha, W. J. Cho and J. W. Park (2013). "Tumor suppressor p53 plays a key role in induction of both tristetraprolin and let-7 in human cancer cells." Nucleic Acids Res **41**(11): 5614-5625.

Lehmann, S. M., C. Kruger, B. Park, K. Derkow, K. Rosenberger, J. Baumgart, T. Trimbuch, G. Eom, M. Hinz, D. Kaul, P. Habbel, R. Kalin, E. Franzoni, A. Rybak, D. Nguyen, R. Veh, O. Ninnemann, O. Peters, R. Nitsch, F. L. Heppner, D. Golenbock, E. Schott, H. L. Ploegh, F. G. Wulczyn and S. Lehnardt (2012). "An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration." Nat Neurosci **15**(6): 827-835.

Leoni, F., G. Fossati, E. C. Lewis, J. K. Lee, G. Porro, P. Pagani, D. Modena, M. L. Moras, P. Pozzi, L. L. Reznikov, B. Siegmund, G. Fantuzzi, C. A. Dinarello and P. Mascagni (2005). "The histone deacetylase inhibitor

ITF2357 reduces production of pro-inflammatory cytokines in vitro and systemic inflammation in vivo." Molecular medicine **11**(1-12): 1-15.

Leslie, N. R., D. Bennett, Y. E. Lindsay, H. Stewart, A. Gray and C. P. Downes (2003). "Redox regulation of PI 3 - kinase signalling via inactivation of PTEN." The EMBO Journal **22**(20): 5501-5510.

Leung, A. K., S. Vyas, J. E. Rood, A. Bhutkar, P. A. Sharp and P. Chang (2011). "Poly(ADP-ribose) regulates stress responses and microRNA activity in the cytoplasm." Mol Cell **42**(4): 489-499.

Leung, A. K. L., J. M. Calabrese and P. A. Sharp (2006). "Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules." Proceedings of the National Academy of Sciences **103**(48): 18125-18130.

Lewis, B. P., C. B. Burge and D. P. Bartel (2005). "Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets." Cell **120**(1): 15-20.

Lewis, D. F. V. (2002). "Oxidative stress: the role of cytochromes P450 in oxygen activation." Journal of Chemical Technology and Biotechnology **77**: 1095-1100.

Li, J. and X. Lu (2013). "The emerging roles of 3' untranslated regions in cancer." Cancer Letters **337**(1): 22-25.

Li, J., K. Wang, X. Chen, H. Meng, M. Song, Y. Wang, X. Xu and Y. Bai (2012). "Transcriptional activation of microRNA-34a by NF-kappa B in human esophageal cancer cells." BMC Molecular Biology **13**(1): 4.

Li, W. and A.-N. Kong (2009). "Molecular mechanisms of Nrf2-mediated antioxidant response." Molecular Carcinogenesis **48**(2): 91-104.

Li, W., N. Thakor, E. Y. Xu, Y. Huang, C. Chen, R. Yu, M. Holcik and A.-N. Kong (2010). "An internal ribosomal entry site mediates redox-sensitive translation of Nrf2." Nucleic Acids Research **38**(3): 778-788.

Li, Y., T. G. VandenBoom, Z. Wang, D. Kong, S. Ali, P. A. Philip and F. H. Sarkar (2010). "miR-146a Suppresses Invasion of Pancreatic Cancer Cells." Cancer Research **70**(4): 1486-1495.

Li, Y., T. Zhang, H. Korkaya, S. Liu, H. F. Lee, B. Newman, Y. Yu, S. G. Clouthier, S. J. Schwartz, M. S. Wicha and D. Sun (2010). "Sulforaphane, a dietary component of broccoli/broccoli sprouts, inhibits breast cancer stem cells." Clin Cancer Res **16**(9): 2580-2590.

- Li, Y., Z. Zhao, C. Xu, Z. Zhou, Z. Zhu and T. You (2014). "HMGA2 induces transcription factor Slug expression to promote epithelial-to-mesenchymal transition and contributes to colon cancer progression." Cancer Letters **355**(1): 130-140.
- Liang, W. C., W. M. Fu, C. W. Wong, Y. Wang, W. M. Wang, G. X. Hu, L. Zhang, L. J. Xiao, D. C. Wan, J. F. Zhang and M. M. Waye (2015). "The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer." Oncotarget **6**(26): 22513-22525.
- Lin, P.-L., J. T. Chang, D.-W. Wu, C.-C. Huang and H. Lee (2016). "Cytoplasmic localization of Nrf2 promotes colorectal cancer with more aggressive tumors via upregulation of PSMD4." Free Radical Biology and Medicine **95**: 121-132.
- Lin, S. L., A. Chiang, D. Chang and S. Y. Ying (2008). "Loss of mir-146a function in hormone-refractory prostate cancer." RNA **14**(3): 417-424.
- Liu, J., F. V. Rivas, J. Wohlschlegel, J. R. Yates, R. Parker and G. J. Hannon (2005). "A role for the P-body component GW182 in microRNA function." Nat Cell Biol **7**(12): 1261-1266.
- Loboda, A., A. Jozkowicz and J. Dulak (2012). "HIF-1 versus HIF-2 — Is one more important than the other?" Vascular Pharmacology **56**(5–6): 245-251.
- Locatelli, F., B. Canaud, K. U. Eckardt, P. Stenvinkel, C. Wanner and C. Zoccali (2003). "Oxidative stress in end - stage renal disease: an emerging threat to patient outcome." Nephrology Dialysis Transplantation **18**(7): 1272-1280.
- Low, S. C., J. W. Harney and M. J. Berry (1995). "Cloning and Functional Characterization of Human Selenophosphate Synthetase, an Essential Component of Selenoprotein Synthesis." Journal of Biological Chemistry **270**(37): 21659-21664.
- Lozano, R., M. Naghavi, K. Foreman, S. Lim, K. Shibuya, V. Aboyans, J. Abraham, T. Adair, R. Aggarwal, S. Y. Ahn, M. A. AlMazroa, M. Alvarado, H. R. Anderson, L. M. Anderson, K. G. Andrews, C. Atkinson, L. M. Baddour, S. Barker-Collo, D. H. Bartels, M. L. Bell, E. J. Benjamin, D. Bennett, K. Bhalla, B. Bikbov, A. B. Abdulhak, G. Birbeck, F. Blyth, I. Bolliger, S. Boufous, C. Bucello, M. Burch, P. Burney, J. Carapetis, H. Chen, D. Chou, S. S. Chugh, L. E. Coffeng, S. D. Colan, S. Colquhoun, K. E. Colson, J. Condon, M. D. Connor, L. T. Cooper, M. Corriere, M. Cortinovis, K. C. de Vaccaro, W. Couser, B. C. Cowie, M. H. Criqui, M. Cross, K. C. Dabhadkar, N. Dahodwala, D. De Leo, L. Degenhardt, A. Delossantos, J. Denenberg, D. C. Des Jarlais, S. D. Dharmaratne, E. R. Dorsey, T. Driscoll, H. Duber, B. Ebel, P. J. Erwin, P. Espindola, M. Ezzati, V. Feigin, A. D. Flaxman, M. H. Forouzanfar, F. G. R. Fowkes, R. Franklin, M. Fransen, M. K. Freeman, S. E. Gabriel, E. Gakidou, F. Gaspari, R. F. Gillum, D. Gonzalez-Medina, Y. A.

Halasa, D. Haring, J. E. Harrison, R. Havmoeller, R. J. Hay, B. Hoen, P. J. Hotez, D. Hoy, K. H. Jacobsen, S. L. James, R. Jasrasaria, S. Jayaraman, N. Johns, G. Karthikeyan, N. Kassebaum, A. Keren, J.-P. Khoo, L. M. Knowlton, O. Kobusingye, A. Koranteng, R. Krishnamurthi, M. Lipnick, S. E. Lipshultz, S. L. Ohno, J. Mabweijano, M. F. MacIntyre, L. Mallinger, L. March, G. B. Marks, R. Marks, A. Matsumori, R. Matzopoulos, B. M. Mayosi, J. H. McAnulty, M. M. McDermott, J. McGrath, Z. A. Memish, G. A. Mensah, T. R. Merriman, C. Michaud, M. Miller, T. R. Miller, C. Mock, A. O. Mocumbi, A. A. Mokdad, A. Moran, K. Mulholland, M. N. Nair, L. Naldi, K. M. V. Narayan, K. Nasser, P. Norman, M. O'Donnell, S. B. Omer, K. Ortblad, R. Osborne, D. Ozgediz, B. Pahari, J. D. Pandian, A. P. Rivero, R. P. Padilla, F. Perez-Ruiz, N. Perico, D. Phillips, K. Pierce, C. A. Pope, E. Porrini, F. Pourmalek, M. Raju, D. Ranganathan, J. T. Rehm, D. B. Rein, G. Remuzzi, F. P. Rivara, T. Roberts, F. R. De León, L. C. Rosenfeld, L. Rushton, R. L. Sacco, J. A. Salomon, U. Sampson, E. Sanman, D. C. Schwebel, M. Segui-Gomez, D. S. Shepard, D. Singh, J. Singleton, K. Sliwa, E. Smith, A. Steer, J. A. Taylor, B. Thomas, I. M. Tleyjeh, J. A. Towbin, T. Truelsen, E. A. Undurraga, N. Venketasubramanian, L. Vijayakumar, T. Vos, G. R. Wagner, M. Wang, W. Wang, K. Watt, M. A. Weinstock, R. Weintraub, J. D. Wilkinson, A. D. Woolf, S. Wulf, P.-H. Yeh, P. Yip, A. Zabetian, Z.-J. Zheng, A. D. Lopez and C. J. L. Murray (2012). "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 **380**(9859): 2095-2128.

Luxen, S., S. A. Belinsky and U. G. Knaus (2008). "Silencing of DUOX NADPH Oxidases by Promoter Hypermethylation in Lung Cancer." Cancer Research **68**(4): 1037-1045.

Ma, L., J. Young, H. Prabhala, E. Pan, P. Mestdagh, D. Muth, J. Teruya-Feldstein, F. Reinhardt, T. T. Onder, S. Valastyan, F. Westermann, F. Speleman, J. Vandesompele and R. A. Weinberg (2010). "miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis." Nat Cell Biol **12**(3): 247-256.

Mah, S. M., C. Buske, R. K. Humphries and F. Kuchenbauer (2010). "miRNA#42: A Passenger Stranded in RNA-Induced Silencing Complex?" **20**(2): 141-148.

Mathers, J. C. and J. E. Hesketh (2007). "The Biological Revolution: Understanding the Impact of SNPs on Diet-Cancer Interrelationships." The Journal of Nutrition **137**(1): 253S-258S.

Mazelin, L., A. Bernet, C. Bonod-Bidaud, L. Pays, S. Arnaud, C. Gespach, D. E. Bredesen, J.-Y. Scoazec and P. Mehlen (2004). "Netrin-1 controls colorectal tumorigenesis by regulating apoptosis." Nature **431**(7004): 80-84.

McBrian, M. A., I. S. Behbahan, R. Ferrari, T. Su, T. W. Huang and K. Li (2013). "Global Histone Acetylation Is Linked to pH." Cancer Discovery **3**(1): 12.

Mercado, N., R. Thimmulappa, C. M. Thomas, P. S. Fenwick, K. K. Chana, L. E. Donnelly, S. Biswal, K. Ito and P. J. Barnes (2011). "Decreased histone deacetylase 2 impairs Nrf2 activation by oxidative stress." Biochemical and Biophysical Research Communications **406**(2): 292-298.

miRBase. (2014). "miRBase Blog." Retrieved 8 Sep 16, 2016, from <http://www.mirbase.org/blog/>.

Murakami, S. and H. Motohashi (2015). "Roles of NRF2 in cell proliferation and differentiation." Free Radic Biol Med.

Murray-Stewart, T., C. L. Hanigan, P. M. Woster, L. J. Marton and R. A. Casero, Jr. (2013). "Histone deacetylase inhibition overcomes drug resistance through a miRNA-dependent mechanism." Mol Cancer Ther **12**(10): 2088-2099.

Myzak, M. C., K. Hardin, R. Wang, R. H. Dashwood and E. Ho (2006). "Sulforaphane inhibits histone deacetylase activity in BPH-1, LnCaP and PC-3 prostate epithelial cells." Carcinogenesis **27**(4): 811-819.

Neilsen, C. T., G. J. Goodall and C. P. Bracken (2012). "IsomiRs – the overlooked repertoire in the dynamic microRNAome." Trends in Genetics **28**(11): 544-549.

Nian, H., B. Delage, E. Ho and R. H. Dashwood (2009). "Modulation of histone deacetylase activity by dietary isothiocyanates and allyl sulfides: studies with sulforaphane and garlic organosulfur compounds." Environmental and molecular mutagenesis **50**(3): 213-221.

Nishida, N., T. Yokobori, K. Mimori, T. Sudo, F. Tanaka, K. Shibata, H. Ishii, Y. Doki, H. Kuwano and M. Mori (2011). "MicroRNA miR-125b is a prognostic marker in human colorectal cancer." Int J Oncol **38**(5): 1437-1443.

Ohta, T., K. Iijima, M. Miyamoto, I. Nakahara, H. Tanaka, M. Ohtsuji, T. Suzuki, A. Kobayashi, J. Yokota, T. Sakiyama, T. Shibata, M. Yamamoto and S. Hirohashi (2008). "Loss of Keap1 Function Activates Nrf2 and Provides Advantages for Lung Cancer Cell Growth." Cancer Research **68**(5): 1303-1309.

Onder, T. T., P. B. Gupta, S. A. Mani, J. Yang, E. S. Lander and R. A. Weinberg (2008). "Loss of E-Cadherin Promotes Metastasis via Multiple Downstream Transcriptional Pathways." Cancer Research **68**(10): 3645-3654.

Ørom, U. A., F. C. Nielsen and A. H. Lund (2008). "MicroRNA-10a Binds the 5' UTR of Ribosomal Protein mRNAs and Enhances Their Translation." Molecular Cell **30**(4): 460-471.

Pappa, G., H. Bartsch and C. Gerhäuser (2007). "Biphasic modulation of cell proliferation by sulforaphane at physiologically relevant exposure times in a human colon cancer cell line." Molecular Nutrition & Food Research **51**(8): 977-984.

Petrova, T. V., A. Nykänen, C. Norrmén, K. I. Ivanov, L. C. Andersson, C. Haglund, P. Puolakkainen, F. Wempe, H. von Melchner, G. Gradwohl, S. Vanharanta, L. A. Aaltonen, J. Saharinen, M. Gentile, A. Clarke, J. Taipale, G. Oliver and K. Alitalo (2008). "Transcription Factor PROX1 Induces Colon Cancer Progression by Promoting the Transition from Benign to Highly Dysplastic Phenotype." Cancer Cell **13**(5): 407-419.

Pi, J., Q. Zhang, J. Fu, C. G. Woods, Y. Hou, B. E. Corkey, S. Collins and M. E. Andersen (2010). "ROS signaling, oxidative stress and Nrf2 in pancreatic beta-cell function." Toxicology and Applied Pharmacology **244**(1): 77-83.

Poynter, M. E. and R. A. Daynes (1998). "Peroxisome Proliferator-activated Receptor α Activation Modulates Cellular Redox Status, Represses Nuclear Factor- κ B Signaling, and Reduces Inflammatory Cytokine Production in Aging." Journal of Biological Chemistry **273**(49): 32833-32841.

Prasad, A. K. and P. C. Mishra (2015). "Mechanism of Action of Sulforaphane as a Superoxide Radical Anion and Hydrogen Peroxide Scavenger by Double Hydrogen Transfer: A Model for Iron Superoxide Dismutase." The Journal of Physical Chemistry B **119**(25): 7825-7836.

Rachakonda, G., K. R. Sekhar, D. Jowhar, P. C. Samson, J. P. Wikswo, R. D. Beauchamp, P. K. Datta and M. L. Freeman (2010). "Increased cell migration and plasticity in Nrf2-deficient cancer cell lines." Oncogene **29**(25): 3703-3714.

Rahman, I., J. Marwick and P. Kirkham (2004). "Redox modulation of chromatin remodeling: impact on histone acetylation and deacetylation, NF- κ B and pro-inflammatory gene expression." Biochemical pharmacology **68**(6): 1255-1267.

Rajendran, P., A. I. Kidane, T.-W. Yu, W.-M. Dashwood, W. H. Bisson, C. V. Löhr, E. Ho, D. E. Williams and R. H. Dashwood (2013). "HDAC turnover, CtIP acetylation and dysregulated DNA damage signaling in colon cancer cells treated with sulforaphane and related dietary isothiocyanates." Epigenetics **8**(6): 577-576.

Ramos-Gomez, M., M.-K. Kwak, P. M. Dolan, K. Itoh, M. Yamamoto, P. Talalay and T. W. Kensler (2001). "Sensitivity to carcinogenesis is increased and chemoprotective efficacy of enzyme inducers is lost in nrf2 transcription factor-deficient mice." Proceedings of the National Academy of Sciences **98**(6): 3410-3415.

Reichard, J. F., G. T. Motz and A. Puga (2007). "Heme oxygenase-1 induction by NRF2 requires inactivation of the transcriptional repressor BACH1." Nucleic Acids Research **35**(21): 7074-7086.

Reid, G., S. C. Kao, N. Pavlakis, H. Brahmabhatt, J. MacDiarmid, S. Clarke, M. Boyer and N. van Zandwijk (2016). "Clinical development of TargomiRs, a miRNA mimic-based treatment for patients with recurrent thoracic cancer." Epigenomics **8**(8): 1079-1085.

Rodrigues, N. R., A. Rowan, M. E. Smith, I. B. Kerr, W. F. Bodmer, J. V. Gannon and D. P. Lane (1990). "p53 mutations in colorectal cancer." Proc Natl Acad Sci U S A **87**(19): 7555-7559.

Sachdeva, M., S. Zhu, F. Wu, H. Wu, V. Walia, S. Kumar, R. Elble, K. Watabe and Y. Y. Mo (2009). "p53 represses c-Myc through induction of the tumor suppressor miR-145." Proceedings of the National Academy of Sciences of the United States of America **106**(9): 3207-3212.

Saddawi-Konefka, R., R. Seelige, E. T. Gross, E. Levy, S. C. Searles, A. Washington, Jr., E. K. Santosa, B. Liu, T. E. O'Sullivan, O. Harismendy and J. D. Bui (2016). "Nrf2 Induces IL-17D to Mediate Tumor and Virus Surveillance." Cell Rep.

Saleh, A. D., J. E. Savage, L. Cao, B. P. Soule, D. Ly, W. DeGraff, C. C. Harris, J. B. Mitchell and N. L. Simone (2011). "Cellular Stress Induced Alterations in MicroRNA let-7a and let-7b Expression Are Dependent on p53." PLoS ONE **6**(10): e24429.

Salvi, A., C. Sabelli, S. Moncini, M. Venturin, B. Arici, P. Riva, N. Portolani, S. M. Giuliani, G. De Petro and S. Barlati (2009). "MicroRNA-23b mediates urokinase and c-met downmodulation and a decreased migration of human hepatocellular carcinoma cells." FEBS Journal **276**(11): 2966-2982.

Schell, John C., Kristofor A. Olson, L. Jiang, Amy J. Hawkins, Jonathan G. Van Vranken, J. Xie, Robert A. Egnatchik, Espen G. Earl, Ralph J. DeBerardinis and J. Rutter (2014). "A Role for the Mitochondrial Pyruvate Carrier as a Repressor of the Warburg Effect and Colon Cancer Cell Growth." Molecular Cell **56**(3): 400-413.

Schmittgen, T. D. (2008). "Regulation of microRNA processing in development, differentiation and cancer." Journal of Cellular and Molecular Medicine **12**(5b): 1811-1819.

Sestili, P., M. Paolillo, M. Lenzi, E. Colombo, L. Vallorani, L. Casadei, C. Martinelli and C. Fimognari (2010). "Sulforaphane induces DNA single strand breaks in cultured human cells." Mutat Res **689**(1-2): 65-73.

Shah, N. M., S. A. Rushworth, M. Y. Murray, K. M. Bowles and D. J. MacEwan (2013). "Understanding the role of NRF2-regulated miRNAs in human malignancies." Oncotarget **4**(8): 1130-1142.

Shah, N. M., L. Zaitseva, K. M. Bowles, D. J. MacEwan and S. A. Rushworth (2015). "NRF2-driven miR-125B1 and miR-29B1 transcriptional regulation controls a novel anti-apoptotic miRNA regulatory network for AML survival." Cell Death Differ **22**(4): 654-664.

Shan, Y., L. Zhang, Y. Bao, B. Li, C. He, M. Gao, X. Feng, W. Xu, X. Zhang and S. Wang (2013). "Epithelial-mesenchymal transition, a novel target of sulforaphane via COX-2/MMP2, 9/Snail, ZEB1 and miR-200c/ZEB1 pathways in human bladder cancer cells." The Journal of Nutritional Biochemistry **24**(6): 1062-1069.

Sharma, R., A. Sharma, P. Chaudhary, M. Sahu, S. Jaiswal, S. Awasthi and Y. C. Awasthi (2012). "Role of 4-hydroxynonenal in chemopreventive activities of sulforaphane." Free Radic Biol Med **52**(11-12): 2177-2185.

Simone, N. L., B. P. Soule, D. Ly, A. D. Saleh, J. E. Savage, W. DeGraff, J. Cook, C. C. Harris and D. Gius (2009). "Ionizing Radiation-Induced Oxidative Stress Alters miRNA Expression " PLOS One **4**(7): e6377.

Singh, A., C. Happel, S. K. Manna, G. Acquaaah-Mensah, J. Carrerero, S. Kumar, P. Nasipuri, K. W. Krausz, N. Wakabayashi, R. Dewi, L. G. Boros, F. J. Gonzalez, E. Gabrielson, K. K. Wong, G. Girnun and S. Biswal (2013). "Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis." J Clin Invest **123**(7): 2921-2934.

Singh, K., S. L. Connors, E. A. Macklin, K. D. Smith, J. W. Fahey, P. Talalay and A. W. Zimmerman (2014). "Sulforaphane treatment of autism spectrum disorder (ASD)." Proceedings of the National Academy of Sciences **111**(43): 15550-15555.

Slaby, O., M. Sachlova, V. Brezkova, R. Hezova, A. Kovarikova, S. Bischofova, S. Sevcikova, J. Bienertova-Vasku, A. Vasku, M. Svoboda and R. Vyzula (2013). "Identification of MicroRNAs Regulated by Isothiocyanates and Association of Polymorphisms Inside Their Target Sites with Risk of Sporadic Colorectal Cancer." Nutrition and cancer **65**(2): 247-254.

Slaby, O., M. Svoboda, J. Michalek and R. Vyzula (2009). "MicroRNAs in colorectal cancer: translation of molecular biology into clinical application." Mol Cancer **8**: 102.

Smith, D., M. Ballal, R. Hodder, G. Soin, S. N. Selvachandran and D. Cade (2006). "Symptomatic presentation of early colorectal cancer." Annals of the Royal College of Surgeons of England **88**(2): 185-190.

Sorefan, K., H. Pais, A. E. Hall, A. Kozomara, S. Griffiths-Jones, V. Moulton and T. Dalmay (2012). "Reducing ligation bias of small RNAs in libraries for next generation sequencing." Silence **3**(1): 1-11.

Spira, A., J. E. Beane, V. Shah, K. Steiling, G. Liu, F. Schembri, S. Gilman, Y.-M. Dumas, P. Calner, P. Sebastiani, S. Sridhar, J. Beamis, C. Lamb, T. Anderson, N. Gerry, J. Keane, M. E. Lenburg and J. S. Brody (2007). "Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer." Nat Med **13**(3): 361-366.

Sriuranpong, V., M. W. Borges, R. K. Ravi, D. R. Arnold, B. D. Nelkin, S. B. Baylin and D. W. Ball (2001). "Notch signaling induces cell cycle arrest in small cell lung cancer cells." Cancer Res **61**(7): 3200-3205.

Subramanian, M., S. R. Rao, P. Thacker, S. Chatterjee and D. Karunagaran (2014). "MiR-29b downregulates canonical Wnt signaling by suppressing coactivators of beta-catenin in human colorectal cancer cells." J Cell Biochem **115**(11): 1974-1984.

Tan, G. C., E. Chan, A. Molnar, R. Sarkar, D. Alexieva, I. M. Isa, S. Robinson, S. Zhang, P. Ellis, C. F. Langford, P. V. Guillot, A. Chandrashekran, N. M. Fisk, L. Castellano, G. Meister, R. M. Winston, W. Cui, D. Baulcombe and N. J. Dibb (2014). "5' isomiR variation is of functional and evolutionary importance." Nucleic Acids Res **42**(14): 9424-9435.

Taniguchi, K., N. Sugito, M. Kumazaki, H. Shinohara, N. Yamada, N. Matsushashi, M. Futamura, Y. Ito, Y. Otsuki, K. Yoshida, K. Uchiyama and Y. Akao (2015). "Positive feedback of DDX6/c-Myc/PTB1 regulated by miR-124 contributes to maintenance of the Warburg effect in colon cancer cells." Biochim Biophys Acta **1852**(9): 1971-1980.

Tierens, K. F. M.-J., B. P. H. J. Thomma, M. Brouwer, J. Schmidt, K. Kistner, A. Porzel, B. Mauch-Mani, B. P. A. Cammue and W. F. Broekaert (2001). "Study of the Role of Antimicrobial Glucosinolate-Derived Isothiocyanates in Resistance of Arabidopsis to Microbial Pathogens." Plant Physiology **125**(4): 1688-1699.

Tili, E., J.-J. Michaille, D. Wernicke, H. Alder, S. Costinean, S. Volinia and C. M. Croce (2011). "Mutator activity induced by microRNA-155 (miR-155) links inflammation and cancer." Proceedings of the National Academy of Sciences **108**(12): 4908-4913.

Trachootham, D., Y. Zhou, H. Zhang, Y. Demizu, Z. Chen, H. Pelicano, P. J. Chiao, G. Achanta, R. B. Arlinghaus, J. Liu and P. Huang (2006). "Selective killing of oncogenically transformed cells through a ROS-mediated mechanism by beta-phenylethyl isothiocyanate." Cancer cell **10**(3): 241-252.

Tse, G. and G. D. Eslick (2014). "Cruciferous vegetables and risk of colorectal neoplasms: a systematic review and meta-analysis." Nutr Cancer **66**(1): 128-139.

USFDA. (2005). "Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers." Guidance for Industry Retrieved 7 Sep 16, 2016, from <http://www.fda.gov/downloads/Drugs/Guidances/UCM078932.pdf>.

Vasudevan, S., Y. Tong and J. A. Steitz (2007). "Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation." Science **318**(5858): 1931-1934.

Wang, B., W. Li, H. Liu, L. Yang, Q. Liao, S. Cui, H. Wang and L. Zhao (2014). "miR-29b suppresses tumor growth and metastasis in colorectal cancer via downregulating Tiam1 expression and inhibiting epithelial-mesenchymal transition." Cell Death Dis **5**: e1335.

Wang, F., P. Zhang, Y. Ma, J. Yang, M. P. Moyer, C. Shi, J. Peng and H. Qin (2012). "NIRF is frequently upregulated in colorectal cancer and its oncogenicity can be suppressed by let-7a microRNA." Cancer Letters **314**(2): 223-231.

Wang, S., H. Yu and J. K. Wickliffe (2011). "Limitation of the MTT and XTT assays for measuring cell viability due to superoxide formation induced by nano-scale TiO₂." Toxicol In Vitro **25**(8): 2147-2151.

Westholm, J. O. and E. C. Lai (2011). "Mirtrons: microRNA biogenesis via splicing." Biochimie **93**(11): 1897-1904.

Wiesen, J. L. and T. B. Tomasi (2009). "Dicer is regulated by cellular stresses and interferons." Molecular Immunology **46**(6): 1222-1228.

Wilczynska, A. and M. Bushell (2015). "The complexity of miRNA-mediated repression." Cell Death Differ **22**(1): 22-33.

World Cancer Research Fund International. (2015). "Colorectal cancer statistics." Retrieved 7 Sep 16, 2016, from <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/colorectal-cancer-statistics>.

World Health Organization: International Agency for Research on Cancer. (2016). "Colorectal Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012." Retrieved 17 Oct 2016, from http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.

Wu, C.-C., F.-M. Tsai, R.-Y. Shyu, Y.-M. Tsai, C.-H. Wang and S.-Y. Jiang (2011). "G protein-coupled receptor kinase 5 mediates Tazarotene-induced

gene 1-induced growth suppression of human colon cancer cells." BMC Cancer **11**(1): 175.

Wyman, S. K., E. C. Knouf, R. K. Parkin, B. R. Fritz, D. W. Lin, L. M. Dennis, M. A. Krouse, P. J. Webster and M. Tewari (2011). "Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity." Genome Res **21**(9): 1450-1461.

Xu, P., M. Billmeier, I. Mohorianu, D. Green, W. D. Fraser and T. Dalmay (2015). "An improved protocol for small RNA library construction using High Definition adapters." Methods in Next Generation Sequencing **2**(1).

Xu, S., R. Zhang, J. Niu, D. Cui, B. Xie, B. Zhang, K. Lu, W. Yu, X. Wang and Q. Zhang (2012). "Oxidative Stress Mediated-Alterations of the MicroRNA Expression Profile in Mouse Hippocampal Neurons." International journal of molecular sciences **13**(12): 16945-16960.

Xu, X. M., B. A. Carlson, H. Mix, Y. Zhang, K. Saira, R. S. Glass, M. J. Berry, V. N. Gladyshev and D. L. Hatfield (2007). "Biosynthesis of selenocysteine on its tRNA in eukaryotes." PLoS Biol **5**(1): e4.

Xue, M., H. Momiji, N. Rabbani, T. Bretschneider, D. A. Rand and P. J. Thornalley (2015). "Frequency modulated translocational oscillations of Nrf2, a transcription factor functioning like a wireless sensor." Biochem Soc Trans **43**(4): 669-673.

Yamagishi, H., H. Kuroda, Y. Imai and H. Hiraishi (2016). "Molecular pathogenesis of sporadic colorectal cancers." Chinese Journal of Cancer **35**(1): 1-8.

Yamamoto, Y., Y. Yoshioka, K. Minoura, R.-u. Takahashi, F. Takeshita, T. Taya, R. Horii, Y. Fukuoka, T. Kato, N. Kosaka and T. Ochiya (2011). "An integrative genomic analysis revealed the relevance of microRNA and gene expression for drug-resistance in human breast cancer cells." Molecular Cancer **10**(1): 135.

Yan, B., Q. Guo, F. J. Fu, Z. Wang, Z. Yin, Y. B. Wei and J. R. Yang (2015). "The role of miR-29b in cancer: regulation, function, and signaling." Onco Targets Ther **8**: 539-548.

Ye, L., A. T. Dinkova-Kostova, K. L. Wade, Y. Zhang, T. A. Shapiro and P. Talalay (2002). "Quantitative determination of dithiocarbamates in human plasma, serum, erythrocytes and urine: pharmacokinetics of broccoli sprout isothiocyanates in humans." Clinica Chimica Acta **316**(1-2): 43-53.

Yun, J., Y. N. Jo, S.-K. Park, J.-A. Kim, J. Yang, C. W. Lee, W. K. Yoon, O.-J. Lee, S.-B. Han, R. C. Doebele and J. S. Kang (2013). "Abstract 3878:

Bach1 promotes liver metastasis of colorectal cancer cells by regulating c-Myc and SOX4." Cancer Research **73**(8 Supplement): 3878.

Zanichelli, F., S. Capasso, G. Bernardo, M. Cipollaro, E. Pagnotta, M. Carteni, F. Casale, R. Iori, A. Giordano and U. Galderisi (2012). "Low concentrations of isothiocyanates protect mesenchymal stem cells from oxidative injuries, while high concentrations exacerbate DNA damage." Apoptosis **17**(9): 964-974.

Zawistowski, J. S., K. Nakamura, J. S. Parker, D. A. Granger, B. T. Golitz and G. L. Johnson (2013). "MicroRNA 9-3p Targets β 1 Integrin To Sensitize Claudin-Low Breast Cancer Cells to MEK Inhibition." Molecular and Cellular Biology **33**(11): 2260-2274.

Zhang, Q., J. Pi, C. G. Woods and M. E. Andersen (2010). "A systems biology perspective on Nrf2-mediated antioxidant response." Toxicol Appl Pharmacol **244**(1): 84-97.

Zhang, X., X. Chen, J. Lin, T. Lwin, G. Wright, L. C. Moscinski, W. S. Dalton, E. Seto, K. Wright, E. Sotomayor and J. Tao (2012). "Myc represses miR-15a/miR-16-1 expression through recruitment of HDAC3 in mantle cell and other non-Hodgkin B-cell lymphomas." Oncogene **31**(24): 3002-3008.

Zhang, Y. (2000). "Role of glutathione in the accumulation of anticarcinogenic isothiocyanates and their glutathione conjugates by murine hepatoma cells." Carcinogenesis **21**(6): 1175-1182.

Zhang, Y. (2000). "Role of glutathione in the accumulation of anticarcinogenic isothiocyanates and their glutathione conjugates by murine hepatoma cells." Carcinogenesis **21**(6): 1175-1182.

Zhao, S. and M.-F. Liu (2009). "Mechanisms of microRNA-mediated gene regulation." Science in China Series C: Life Sciences **52**(12): 1111-1116.